

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:35:30 ; Search time 1314.21 Seconds

(without alignments)
10739.422 Million cell updates/sec

Title: US-09-614-150A-2

Sequence: 1 atcaggaagcagagacaga.....ttcttaacataaagggttag 345

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_ph:*
24: em_pat:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrc:*
38: em_ey:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	325.2	94.3	35353	2	AC020055	AC020055 Drosophil
2	325.2	94.3	181141	3	AC008346	AC008346 Drosophil
3	325.2	94.3	185668	3	AC008097	AC008097 Drosophil
4	325.2	94.3	281797	3	AB003794	AB003794 Drosophil
5	80.6	23.4	14637	3	AF141921	AF141921 Drosophil
6	80.6	23.4	15816	2	AC020002	AC020002 Drosophil
7	80.6	23.4	175506	3	AC008004	AC008004 Drosophil
8	80.6	23.4	290783	3	AE003803	AE003803 Drosophil
9	57.6	16.7	504	3	AF141920	AF141920 Drosophil
10	57.6	16.7	541	3	AY071587	AY071587 Drosophil
11	54	15.7	615	6	BD027650	BD027650 Sequence
12	52.8	15.3	291	9	AF178431	AF178431 Homo sapi
13	52.8	15.3	452	6	BD030271	BD030271 Sequence
14	52.8	15.3	603	6	BD111742	BD111742 EST and e
15	52.8	15.3	670	9	AF165516	AF165516 Homo sapi
16	52.8	15.3	651	9	AK026864	AK026864 Homo sapi
17	52.8	15.3	670	9	AF165516	AF165516 Homo sapi
18	52.8	15.3	702	9	AF132750	AF132750 Homo sapi
19	52.8	15.3	728	9	BC002481	BC002481 Homo sapi
20	52.8	15.3	740	6	AX015354	AX015354 Sequence
21	52.8	15.3	765	6	AX230452	AX230452 Sequence
22	52.4	15.2	396	6	AX093234	AX093234 Sequence
23	50.8	14.7	85539	2	AC019877	AC019877 Drosophil
24	50.8	14.7	85567	3	AC004276	AC004276 Drosophil
25	50.8	14.7	184650	3	AC092186	AC092186 Drosophil
26	50.8	14.7	303641	3	AE003583	AE003583 Drosophil
27	49	14.2	14637	3	AF141921	AF141921 Drosophil
28	49	14.2	15816	2	AC020002	AC020002 Drosophil
29	49	14.2	175506	3	AC008004	AC008004 Drosophil
30	49	14.2	290783	3	AE003803	AE003803 Drosophil
31	48	13.9	668	5	BC046084	BC046084 Danio rer
32	48	13.9	672	10	AF073839	AF073839 Rattus no
33	48	13.9	742	10	AY026512	AY026512 Rattus no
34	48	13.9	749	6	AX230453	AX230453 Sequence
35	48	13.9	2823	6	AX230454	AX230454 Sequence
36	46.4	13.4	953	10	BC029172	BC029172 Mus muscu
37	43.4	12.6	155345	9	AC137627	AC137627 Homo sapi
38	43.4	12.6	177147	9	AC007539	AC007539 Homo sapi
39	43.2	12.5	9710	2	AC014687	AC014687 Drosophil
40	43.2	12.5	172748	3	AC018484	AC018484 Drosophil
41	43.2	12.5	186695	3	AC013351	AC013351 Drosophil
42	43.2	12.5	298116	3	AE003782	AE003782 Drosophil
43	40.4	11.7	566	3	AB074929	AB074929 Clona int
44	38.8	11.2	112006	2	AC020467	AC020467 Drosophil
45	38.8	11.2	149592	2	AC005718	AC005718 Drosophil

ALIGNMENTS

RESULT 1
AC020055 35353 bp DNA linear HTG 03-JAN-2000
LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
AC020055
DEFINITION AC020055 GI:6664842
AC020055
VERSION AC020055.1
KEYWORDS HTG; HTGS PHASR2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 35353)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission

Pred. No. is the number of results predicted by chance to have a

ACCESSION BACR20D06, complete sequence.
VERSION AC008097
KEYWORDS HG. GI:14327744
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
REFERENCE 1 (bases 1 to 185668)
AUTHORS Gelniker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busan,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Frise,E., Galle,R.F., Gary,N.S., George,R.A., Gonzalez,M., Houch,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phoumenavong,S., Piltman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Swirskae,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
TITLE Sequencing of Drosophila chromosome 2R, region 56F-56F
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 185668)
AUTHORS Gelniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Bienenhoff,C., Champe,M., Chavez,C., Chew,M., Cieciolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,K.N.L., Hoskins,R.A., Houston,K.A., Humast,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomutan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Shtir,E., Swirskae,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.
TITLE Direct Submission
JOURNAL Submitted (22-JUL-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT On Jun 7, 2001 this sequence version replaced gi:6850230.
Sequence submitted by:
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.
FEATURES
location/Qualifiers
source
1..185668
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2R"
/map="56F-56F"
/wgs="56F-56F"
/clone="BACR20D06 (D943)"
/clone_id="RPCI-98 (Rosewell) Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
PBAC3.6")
BASE COUNT 53603 a 40080 c 40061 g 51924 t
ORIGIN
Query Match 94.3%; Score 325.2; DB 3; Length 185668;
Best Local Similarity 99.1%; Pred. No. 3.9e-92;
Matches 327; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 23518 CCCGGCGTGGAGGACATATTGATCATGAATCACTCCGGTGTGCGGTGAACCTCGANG 23577
136 GATGTCAGAGAGGCGTTGACATACGCTGTCTATATGCAATTTGGGGAGAGTCCGAG 195
23578 GATGTCAGAGAGGCGTTGACATACGCTGTCTATATGCAATTTGGGGAGAGTCCGAG 23637
196 GCGTTCCTCTCCAAATGAGAGCCAGCCCAAAATTTGACTCTACTAGAGTTCGACCAAG 255
23638 GCGTTCCTCTCCAAATGAGAGCCAGCCCAAAATTTGACTCTACTAGAGTTCGACCAAG 23697
Db 256 TATCAGAGAGTGTCTATTACCAATGCGCAAGATCACGTTTGGTGTTCGAATGCC 315
23698 TATCAGAGAGTGTCTATTACCAATGCGCAAGATCACGTTTGGTGTTCGAATGCC 23757
Qy 316 AAGATACCTTTTCTTAACATAAAGGTTAG 345
Db 23758 AAGATACCTTTTCTTAACATAAAGGTTAG 23787
RESULT 4
LOCUS AE003794
DEFINITION Drosophila melanogaster chromosome 2R, section 56 of 74 of the complete sequence.
ACCESSION AE003794.3 GI:21626923
VERSION AE003794.3
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
REFERENCE 1 (bases 1 to 281797)
AUTHORS Adams,M.D., Gelniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brannon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,B.G., Heit,G., Nelson,C.R., Gabot,G.L., Abail,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Baller,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brockstein,P., Brotier,P., Burris,K.C., Busan,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablo,S.B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dunn,P., Durbin,K.J., Evansgelistra,C.C., Ferraz,C., Finkov,B.C., Dunn,P., Durbin,K.J., Evansgelistra,C.C., Ferraz,C., Ferreira,S., Fleischmann,W., Foster,K., Gabrielian,A.B., Gary,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Helt,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kienison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lai,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mowbray,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusken,D.R., Pacleb,J.M., Palazzolo,M., Piltman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Swirskae,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A., Weinstein,G.M., Weissbach,J., Williams,S.M., Woodruff, W., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhang,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
TITLE The genome sequence of Drosophila melanogaster
JOURNAL Science 287 (5461), 2185-2195 (2000)

MEDLINE
20196006
PUBMED
10731132
REFERENCE
AUTHORS
Celinker,S.E., Adams,M.D., Krommiller,B., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C., Rogers,Y., Banzon,J., An,H., Baldwin,D., Bazon,J., Beeson,K.Y., Busan,D.A., Carlson,J.W., Center,A., Champs,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Gonzalez,S., Frisoe,E., Galle,R.F., Gary,N.S., George,R.A., Gerzella,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Matzel,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Paclet,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phuaenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Shapiro,M., Strong,R., Swirskas,R., Tector,C., Tyler,D., Williams,S.M., Zavei,J.S., Smith,H.O., Venter,J.C. and Rubin,G.M.
TITLE
JOURNAL
REFERENCE
AUTHORS
Miera,S., Crosby,M.A., Matthews,B.B., Bayraktaroglu,L., Campbell,K., Hradecky,P., Huang,Y., Kaminker,J.S., Prochuk,S.E., Smith,C.D., Tupy,J.L., Bergman,C.M., Bertram,B.P., Carlson,J.W., Celinker,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Frisoe,E., de Grey,A.D.N.J., Harris,N.L., Krommiller,B., Marshall,B., Milburn,G.H., Richter,J., Russo,S., Searle,S.M.J., Smith,E., Shu,S., Smutniak,F., Whitfield,E.J., Ashburner,M., Gelbart,W.M., Rubin,G.M., Mungall,C.J. and Lewis,S.E.
TITLE
JOURNAL
REFERENCE
AUTHORS
FlyBase
Direct Submission
Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA
6 (bases 1 to 281797)
FlyBase
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
5 (bases 1 to 281797)
FlyBase
Direct Submission
Submitted (23-JAN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA
On Jun 28, 2002 this sequence version replaced gi:10727440.
COMMENT
FEATURES
source
1..281797
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="2R"
/note="genotype: Y[2]; cn[1] bw[1] sp[1]; Rh6[1]"
8519..9135
/gene="Obp56e"
/locus_tag="CG8462"
/note="last curated on Tue Feb 19 13:40:49 PST 2002"
/map="56B6-56B6"
/db_xref="FLYBASE:FBgn0034471"
join(8519..8627,8684..9135)
/gene="Obp56e"
/locus_tag="CG8462"
/product="CG8462-RA"
/db_xref="FLYBASE:FBgn0034471"
join(8571..8627,8684..9025)
/gene="Obp56e"
/locus_tag="CG8462"
/note="Obp56e gene product"
/codon_start=1
/product="CG8462-PA"
/protein_id="AAF57519.1"
/db_xref="GI:7302433"

/db_xref="FLYBASE:FBgn0034471"
/translation="MKVFFVPAALALSLASAVGLTDSQAKAKQAKACVKGEGITK
EQAIALRSGNPADSPDKVCFANCFLEQGLVANQGIKPDVVLAKGPIAGEANVKEV
OAKCDSITKADKCDTSTYLYKCYENHAF"
9593..10024
/gene="Obp56f"
/locus_tag="CG30450"
/note="last curated on Wed Sep 25 15:51:14 PDT 2002"
/map="56B6-56B6"
/db_xref="FLYBASE:FBgn004353"
join(9593..9652,9707..10024)
/gene="Obp56f"
/locus_tag="CG30450"
/product="CG30450-RA"
/db_xref="FLYBASE:FBgn004353"
join(9593..9652,9707..10024)
/gene="Obp56f"
/locus_tag="CG30450"
/note="Obp56f gene product"
/codon_start=1
/product="CG30450-PA"
/protein_id="PAM68412.1"
/db_xref="GI:21826924"
/db_xref="FLYBASE:FBgn004353"
/translation="MKVFLFLFISAIWDAQPCMSSEKIKACLKRQLGYTTENTKF
DAKEDLSQKCFYHCLLEKGVANDAISSQPRKVLKKYGYITDDELEKAEKCHS
IKXSGCELGVELIKCYGSITKH"
10250..11047
/locus_tag="CG8517"
/note="last curated on Fri Nov 08 16:27:41 PST 2002"
/map="56B6-56B6"
/db_xref="FLYBASE:FBgn0034472"
join(10250..10520,10582..11047)
/locus_tag="CG8517"
/product="CG8517-RA"
/db_xref="FLYBASE:FBgn0034472"
join(10348..10520,10582..10846)
/locus_tag="CG8517"
/note="CG8517 gene product"
/codon_start=1
/product="CG8517-PA"
/protein_id="AAF57518.2"
/db_xref="GI:28380711"
/db_xref="FLYBASE:FBgn0034472"
/translation="MSSYRRLLPRAKILMLMKANRGMTSGKOPIDVETRKDFEORV
INSDRIVVDVFAHSCGCKALAPLENVSEQEGVRLLARDVIDHSHLADYNGS
VBSLVITISGKVVNRVGIQTSEYLRKMLHKAVPRPPEDTEN"
11739..11812
/locus_tag="CR30215"
/map="56B6-56B6"
11739..11812
/locus_tag="CR30215"
/product="CRNA-11e"
/db_xref="FLYBASE:FBgn0050215"
12082..19495
/locus_tag="TEI18966"
/map="56B6-56B6"
12082..19495
/transposon="b1ood[856"
/db_xref="FLYBASE:FBt10018966"
complement(21794..21865)
/locus_tag="CR30452"
/note="synonym: CR30216"
/map="56F1-56F1"
complement(21794..21865)
/locus_tag="CR30452"
/product="CRNA-Met"
/db_xref="FLYBASE:FBgn0050452"
22079..22150
/locus_tag="CR30218"
/map="56F1-56F1"
22079..22150
/locus_tag="CR30218"

/product="tRNA-Met"
/db_xref="FLYBASE:FBgn0050218"
complement (22316..22387)
gene

Query Match 94.3%; Score 325.2; DB 3; Length 281797;
Best Local Similarity 99.1%; Pred. No. 4.2e-92;
Matches 327; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 16 ACAGATCCAAAGGAACTAAAGTTACGTGAAGAGTATTTGCAAGTGCAGAGAAA 75
DB 247689 ACCGAGCCAAAGCACTAAAGTTACGTGAAGAGTATTTGCAAGTGCAGAGAAA 247748
QY 76 CCCGCGGTGAGAGACATATTTGATCATGATCACTCGGGTGTGCGGTGAAGAACTCGATG 135
DB 247749 CCCGCGGTGAGAGACATATTTGATCATGATCACTCGGGTGTGCGGTGAAGAACTCGATG 247808
QY 136 GATCGTCAGAGAGGCTTGCAGTACGCTGTCTATATGACAAATTTGCGGAGAAAGTGCAG 195
DB 247809 GATCGTCAGAGAGGCTTGCAGTACGCTGTCTATATGACAAATTTGCGGAGAAAGTGCAG 247868
QY 196 GCGTTCCTCTCCAAATAGAGCCAGCCAAATTTGACTGAGAGTTCGTACCAAG 255
DB 247869 GCGTTCCTCTCCAAATAGAGCCAGCCAAATTTGACTGAGAGTTCGTACCAAG 247928
QY 256 TATCAGAGAGTGTCTATTAACACAGATCCAGATCAACGTTTGTGTTGAGAAATGCC 315
DB 247929 TATCAGAGAGTGTCTATTAACACAGATCCAGATCAACGTTTGTGTTGAGAAATGCC 247988
QY 316 AAGATACCTTTTCTTAACATTAAGGTTAG 345
DB 247989 AAGATACCTTTTCTTAACATTAAGGTTAG 248018

RESULT 5
AF141921/c 14637 bp DNA linear INV 31-JUL-1999
LOCUS Drosophila melanogaster chromosome 2 map 54B roadblock (robl) gene,
DEFINITION complete cds.
ACCESSION AF141921
VERSION AF141921.1 GI:5668625
KEYWORDS
ORGANISM Drosophila melanogaster (fruit fly)
LOCUS Drosophila melanogaster
DEFINITION Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 14637)
Bowman,A.B., Patel-King,R.S., Benashski,S.E., McCaffery,J.M.,
Goldstein,L.S. and King,S.M.
Drosophila roadblock and Chlamydomonas LC7: a conserved family of
dynein-associated proteins involved in axonal transport, flagellar
motility, and mitosis
JOURNAL J. Cell Biol. 146 (1), 165-180 (1999)
MEDLINE 99333306
PUBMED 10402468
REFERENCE 2 (bases 1 to 14637)
Bowman,A.B. and Goldstein,L.S.B.
Direct Submission
TITLE Submitted (09-APR-1999) Pharmacology, University of California, San
JOURNAL Diego, 9500 Gilman Dr. 0683, La Jolla, CA 92093-0683, USA
FEATURES
source 1. 14637
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="54B"
/clone="BDGP P1 clone DS02859"
7698..8332
/gene="robl"
mRNA join(7698..7769,7833..7911,7980..8332)
/gene="robl"
CDS join(7767..7769,7833..7911,7980..8191)

/gene="robl"
/note="LC7 light chain dynein-associated protein family"
/codon_start=1
/product="roadblock"
/protein_id="MAD45987.1"
/db_xref="GI:5668626"
/translation="MSQVEETLKRIOHKGVGTVIVNNEGIPVKSTLDNTTVOYA
GLMSQLADARSRVRLDPSNDMTFRLVRSKKHWEIWAADKDFILIVIONPTD"
BASE COUNT 4501 a 3063 c 2899 g 4174 t
ORIGIN

Query Match 23.4%; Score 80.6; DB 3; Length 14637;
Best Local Similarity 53.7%; Pred. No. 4.1e-14;
Matches 167; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 16 ACAGATCCAAAGGAACTAAAGTTACGTGAAGAGTATTTGCAAGTGCAGAGAAA 75
DB 7280 AATTAACCACTTGAACCTGATCATATGAGAAAAGCAATTCGACTGTGCAATTCGAAA 7221
QY 76 CCCGCGGTGAGAGACATATTTGATCATGATCACTCGGGTGTGCGGTGAAGAACTCGATG 135
DB 7220 AATTAACCTGAGATGATGATGATCTCAACAGTCCGGACATCCGTGATGATGATGATG 7161
QY 136 GATCGTCAGAGAGGCTTGCAGTACGCTGTCTATATGACAAATTTGCGGAGAAAGTGCAG 195
DB 7160 GATCGTCAGAGAGGCTTGCAGTACGCTGTCTATATGACAAATTTGCGGAGAAAGTGCAG 7101
QY 196 GCGTTCCTCTCCAAATAGAGCCAGCCAAATTTGACTGAGAGTTCGTACCAAG 255
DB 7100 CCGGTATATAGTAATTAATATCCACAGAGCACTAATGCTTGGAATATCCACAAAG 7041
QY 256 TATCAGAGAGTGTCTATTAACACAGATCCAGATCAACGTTTGTGTTGAGAAATGCC 315
DB 7040 ACCAGAGAGTGTCTATTAACACAGATCCAGATCAACGTTTGTGTTGAGAAATGCC 6981
QY 316 AAGATACCTTT 326
DB 6980 CATGATTACTT 6970

RESULT 6
AC020002 15816 bp DNA linear HTG 03-JAN-2000
LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
DEFINITION AC020002
ACCESSION AC020002
VERSION AC020002.1 GI:6664895
KEYWORDS HTG; HTGS PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
LOCUS Drosophila melanogaster
DEFINITION Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 15816)
Adams,M. and Venter,J.C.
Direct Submission
TITLE Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
JOURNAL Rockville, MD, USA
COMMENT This sequence was identified as CDM:10211691 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source 1. 15816
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
BASE COUNT 4514 a 3108 c 3356 g 4838 t
ORIGIN

Query Match 23.4%; Score 80.6; DB 2; Length 15816;
Best Local Similarity 53.7%; Pred. No. 4.2e-14;

Matches 167; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 16 ACAGATCCAAAGGCACTAAAGTTAGCTGAGAGAGATTTTGGCAAGTGCAGAGAAA 75
 Db 10283 AAATAGCCACCTGCACTTACGATATGTGAGAAAAGCATTTCCAGCTGGTCAATGCAAA 10342
 QY 76 CCGGCGCTGGAGAGACATATTGATGATGATCACTCGGGGTGGCGGTGAAAACCTTCGATG 135
 Db 10343 AAACATGTTGGAGATGTAGTATGATCTTCACAGAGTCGGGACATCCGGTATGATGATGATG 10402
 QY 136 GATGTGTAGAGAGGCTTGCAGTACGCTGTCTATATGACAAATTTGGCGGAGAAATGCGCAG 195
 Db 10403 GACCGAGAGAGATGTGTCCAAATTTTCGGGACCTTTTACGCTATTTCCAGAGAAAGCTGAAA 10462
 QY 196 GCGTTCCTCTCCAAATATGAGAGCCAGCCAAATTTGACTCTACTGAGAGTTCGTACCAAG 255
 Db 10463 CCGGCTATGATGATTAATAGATCCACGAGCACTACTAATGCTGAGAAATACGACAGA 10522
 QY 256 TATCAGAGAGTGTCTATTACACCAATGCCAAGATCCGTTTGGTGGTTCAGAAATGCC 315
 Db 10523 ACCAAGAGAGTGTCTCTGTCGTCAGACTCCAAAGATTAACACTACTGATGATGACAGACGG 10582
 QY 316 AAGATACCTT 326
 Db 10583 CATGATTACTT 10593

RESULT 7

AC008004/c

LOCUS Drosophila melanogaster, chromosome 2R, region 54B-54C, BAC clone
 DEFINITION BACR21A09, complete sequence.

AC008004

AC008004.5 GI:14280141

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

175506 bp DNA linear INV 02-JUN-2001
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 175506)
 Cealner,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
 Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
 Roget,Y., An,H., Baldwin,D., Bonzon,J., Beeson,K.Y., Busam,D.A.,
 Carlson,J.W., Center,A., Chapple,M., Davenport,L.B., Dietz,S.M.,
 Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
 Ferreira,S., Frise,E., Galle,R.F., Gary,N.S., George,R.A.,
 Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
 Idegawa,C., Jalili,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
 McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
 Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
 Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
 Stapleton,M., Strong,R., Svitkay,R., Tector,C., Williams,S.M.,
 Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
 Sequencing of Drosophila chromosome 2R, region 54B-54C

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

COMMENT

Submitted (09-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
 Laboratory, MS 64-121, Berkeley, CA 94720, USA
 On Jun 2, 2001 this sequence version replaced gi:5851707.
 Sequence submitted by:
 Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720
 This sequence was assembled using end sequences from a whole genome
 shotgun and from subclones of this BAC and its neighboring clones.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our sequence
 archive web site (<http://www.fruitfly.org/sequence/>) or send email
 to bdg@fruitfly.berkeley.edu.

FEATURES

source

1. 175506
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /strain="y; cn bw sp"
 /db_xref="taxon:7227"
 /chromosome="2R"
 /map="54B-54C"
 /clone="BACR21A09 (D936)"
 /clone_1lb="RPCT-98 (Roswell Park Cancer Institute
 Drosophila melanogaster BAC library, partial EcoRI in
 pBAC3.6)"
 BASE COUNT 49750 a 39368 c 38851 g 47537 t
 ORIGIN

Query Match 23.4%; Score 80.6; DB 3; Length 175506;
 Best Local Similarity 53.7%; Pired. No. 6.6e-14;
 Matches 167; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 16 ACAGATCCAAAGGCACTAAAGTTAGCTGAGAGAGATTTTGGCAAGTGCAGAGAAA 75
 Db 47575 AAATAGCCACCTGCACTTACGATATGTGAGAAAAGCATTTCCAGCTGGTCAATGCAAA 47516
 QY 76 CCGGCGCTGGAGAGACATATTGATGATGATCACTCGGGGTGGCGGTGAAAACCTTCGATG 135
 Db 47515 AAACATGTTGGAGATGTAGTATGATCTTCACAGAGTCGGGACATCCGGTATGATGATGATG 47456
 QY 136 GATGTGTAGAGAGGCTTGCAGTACGCTGTCTATATGACAAATTTGGCGGAGAAATGCGCAG 195
 Db 47455 GACCGAGAGAGATGTGTCCAAATTTTCGGGACCTTTTACGCTATTTGAGAGAACTGAAA 47396
 QY 196 GCGTTCCTCTCCAAATATGAGAGCCAGCCAAATTTGACTCTACTGAGAGTTCGTACCAAG 255
 Db 47395 CCGGCTATGATGATTAATAGATCCACGAGCACTACTAATGCTGAGAAATAGCACAAGA 47336
 QY 256 TATCAGAGAGTGTCTATTACACCAATGCCAAGATCCGTTTGGTGGTTCGAATGCC 315
 Db 47335 ACCAAGAGAGTGTCTCTGTCGTCAGACTCCAAAGATTAACACTACTGATGATGACAGACGG 47276
 QY 316 AAGATACCTT 326
 Db 47275 CATGATTACTT 47265

RESULT 8

AE003803/c

LOCUS Drosophila melanogaster chromosome 2R, section 47 of 74 of the
 DEFINITION complete sequence.

AC003803

AE003803.4 GI:28380747

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Submitted (09-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
 Laboratory, MS 64-121, Berkeley, CA 94720, USA
 On Jun 2, 2001 this sequence version replaced gi:5851707.
 Sequence submitted by:
 Lawrence Berkeley National Laboratory, MS 64-121

TITLE
JOURNAL

Beeson, K.Y., Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S.,
Borkova, D., Botchan, M.R., Bouck, J., Brokstein, P., Bottlinger, P.,
Burtis, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A.,
Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B.,
Davies, P., De Pablos, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I.,
Dieter, S.M., Dodson, K., Doup, L.E., Downes, M., Dugan-Rocha, S.,
Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C.,
Ferreira, S., Fleischmann, W., Foster, C., Gabriellian, A.E., Garg, N.S.,
Gelbart, W.M., Glasser, K., Glodke, A., Gong, F., Gorrell, J.H., Gu, Z.,
Guan, P., Harris, M., Harris, N.L., Harvey, D., Helman, T.J.,
Hernandez, J.R., Houck, J.J., Hostin, D., Houston, K.A., Howland, T.J.,
Wei, M.H., Ibegwam, C., Jalali, M., Kalish, F., Kapren, G.H., Ke, Z.,
Kennison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C.,
Kravtitz, S., Kulp, D., Lai, Z., Lesko, P., Lei, Y., Levitsky, A.A.,
Li, Y., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C.,
McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C.,
Morris, J., Mostrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L.,
Muzyk, D.M., Neilson, D.L., Neilson, D.R., Neilson, K.A., Nixon, K.,
Nusskern, D.R., Paciel, J.M., Palazzolo, M., Pittman, G.S., Pan, S.,
Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K.,
Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kimios, I.,
Simpson, M., Skypski, M.P., Smith, T., Spier, E., Spirdling, A.C.,
Stapleton, M., Strong, R., Sun, E., Switzkas, R., Tector, C., Turner, R.,
Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Weissman, D.A.,
Weinstein, G.M., Weissenbach, J., Williams, S.M., Woodgett,
Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F.,
Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H.,
Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O.,
Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.

The genome sequence of *Drosophila melanogaster*
Science 287 (5461), 2185-2195 (2000)

TITLE	Sequencing of <i>Drosophila melanogaster</i> genome
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 290783)
AUTHORS	Mitra, S., Crosby, M.A., Matthews, B.B., Bayraktaroglu, L., ...

TITLE Annotation of *Drosophila melanogaster* genome
 JOURNAL Unpublished
 REFERENCE 4 (bases 1 to 290783)
 AUTHORS Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
 TITLE Direct Submission
 JOURNAL Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 REFERENCE 5 (bases 1 to 290783)

TITLE Direct Submission
JOURNAL Submitted (06-SEP-2002) University of California Berkeley, 539 Life
REFERENCE Sciences Addition, Berkeley, CA 94720, USA.
AUTHORS 6 (bases 1 to 290783)

CONSRM FlyBase
TITLE Direct Submission
JOURNAL Submitted (23-JUN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA
COMMENT On Feb 14, 2003 this sequence version replaced gi:21645220.
FEATURES Location/Qualifiers
source
1..230783

gene

/note="genotype: y[1]; cn[1] bw[1] ep[1]; Rh6[1]"
61..110377
/locus_tag="CG33197"
/note="Tyronymus: CG10941, CG14477; last curated on Tue Nov
26 10:47:06 PST 2002"
/map="54B1-54B5"
/db_xref="FLYBASE:FBgn0053197"
join(61..719,9918..10571,85602..85772,86637..86813,
95621..95662,96488..96605,107585..107669,108326..110377)
/locus_tag="CG33197"
/product="CG33197-RD"
/db_xref="FLYBASE:FBgn0053197"
join(61..719,9918..10571,85602..85772,86637..86813,
93814..95994)
/locus_tag="CG33197"
/product="CG33197-RB"
/db_xref="FLYBASE:FBgn0053197"
join(61..719,9918..10571,85602..85772,86637..88071)
/locus_tag="CG33197"
/product="CG33197-RA"
/db_xref="FLYBASE:FBgn0053197"
join(61..719,9918..10571,30343..33068)
/locus_tag="CG33197"
/product="CG33197-RC"
/db_xref="FLYBASE:FBgn0053197"
join(10383..10571,85602..85772,86637..86813,95621..95662,
96488..96605,107585..107619)
/locus_tag="CG33197"
/note="CG33197 gene product from transcript CG33197-RD"
/codon_start=1
/product="CG33197-PD"
/protein_id="AAO41357.1"
/db_xref="GI:28380748"
/db_xref="FLYBASE:FBgn0053197"
/translation="MAVYVNNNSILNGDSWLDLEVCREFQRNKKCSRODTECKFAHP
PANEVVGKATCYDSTIKRCNDRDKPEPKTFHPQHKDQLINGRNHLAKNALMQ
OMGALPGQPVISGVPAVAIVNPLVTGLIPANSYSYTTGHLVPALLGSDPVSOLGPV
VPTQVVAQOQKIPSDRIEMDVKTGVSPYNDPQSPGWPFKRPAAEKSGIPIVOPGA
TAYOOLMPPYVSPSTGHPGVPRP"
join(10383..10571,85602..85772,86637..86813,93814..94170)
/locus_tag="CG33197"
/note="CG33197 gene product from transcript CG33197-RB"
/codon_start=1
/product="CG33197-PB"
/protein_id="AAO41358.1"
/db_xref="GI:28380749"
/db_xref="FLYBASE:FBgn0053197"
/translation="MAVYVNNNSILNGDSWLDLEVCREFQRNKKCSRODTECKFAHP
PANEVVGKATCYDSTIKRCNDRDKPEPKTFHPQHKDQLINGRNHLAKNALMQ
OMGALPGQPVISGVPAVAIVNPLVTGLIPANSYSYTTGHLVPALLGSDPVSOLGPV
VPTQVVAQOQKIPSDRIEMDVKTGVSPYNDPQSPGWPFKRPAAEKSGIPIVOPGA
TAYOOLMPPYVSPSTGHPGVPRP"
join(10383..10571,85602..85772,86637..86888)
/locus_tag="CG33197"
/note="CG33197 gene product from transcript CG33197-RA"
/codon_start=1
/product="CG33197-PA"
/protein_id="AAO41359.1"
/db_xref="GI:28380750"
/db_xref="FLYBASE:FBgn0053197"
/translation="MAVYVNNNSILNGDSWLDLEVCREFQRNKKCSRODTECKFAHP

CDS

CDS

CDS

MRNA

MRNA

MRNA

CDS
PANVEONGKVATACYSISIKRCNRDKPCKYFHPPHQIKDOLLINGRNLALKNALMQ
OMGIAPQAPISGVPAVATNPYLTGTGTPANSYSRYTGTGLVPLALCPDVTSQLGPV
VPTQVQAOOKIRSDLEYSFGIMPTTITISRYMFWPKKST1"
join(10383, 10571,30343, .30408)
/locus_tag="CG33197"
/note="CG33197 gene product from transcript CG33197-RC"
/locus_start=1
/codon_start=1
/product="CG33197-PC"
/protein_id="AA041360.1"
/db_xref="GI:28380751"
/db_xref="FLYBASE:FBgn0053197"
/translation="MANVVMNLSLNKSDSRMLQLEVCREFQNRKSRDTECKFAHP
PANVEONGKVATACYSISIKVCTPSLYIDLILNIVSKKFC"
11784, .12123
/locus_tag="TE18991"
11784, .12123
/transposon="jockey(}841"
/db_xref="FLYBASE:FBt10018991"
/gene="mm"
/locus_tag="CG14477"
/note="synonym: CG10941, last curated on Thu Mar 21
15:05:51 PST 2002"
/map="54B4-54B5"
/db_xref="FLYBASE:FBgn0015514"
join(<85602, .85772,86637, .86813,95621, .95662,96488, .96605,
107585, 107669,108326, .110377)
/gene="mm"
/locus_tag="CG14477"
/product="CG14477-RA"
/db_xref="FLYBASE:FBgn0015514"
join(<85602, .85772,86637, .86813,95621, .95662,96488, .96605,
107585, 107619)
/gene="mm"
CDS
Query Match 23.4%; Score 80.6; DB 3; Length 290783;
Best Local Similarity 53.7%; Pred. No. 7,2e-14;
Matches 167; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
QY 16 ACAGATCCAAAGCACTAAAGTTAGTGAAGAAGATTTCCCAAGTGCAGAGAAA 75
190592 AAATAGCCACCTCACTCAATGATGTGAAAAAGCATTCGACCTGGTCAATCGAAA 190533
76 CCCGCGGTGAGAGCATATTGATCATGATCACTCGGGTGTGCGGTGAAAACTCGATG 135
190532 AAAATGTTGAGAGTGTAGTATCTCTCAACGAGTCGGGACATCCGCTGATAGTACATG 190473
QY 136 GATGTCAGAGAGGCTTGACGTACGCTGTCTATATGACAAATTTGCGGAGAAAGTCCAG 195
190472 GACGAGAGGATGTGTCCTCAATTTTCGGGACCTTTTCAGCTATTCGAGAAAGTGGAA 190413
QY 196 GCGTTCTCTCCAAAATGAGAGCCAGCCAAAATTTGACTCTACTGAGAGTTCGTACCAAG 255
190412 CGCGGTATGAGTAAATAGATCCACGAGCACTACTTAATGTGAGAAATAGCAGACAGA 190353
QY 256 TATCAGCAGGTGCTCATTTACAGAGATGCCAAGATCCGTTTGTGTGAGTTCAGAAATGCC 315
190352 ACCAAGAGGTGCTCTCGGTGTCGAGATCCAGATTAACATCTCTGTAGTACAGAACCGG 190293
QY 316 AAAGATCTTT 326
190292 CATGATTACTT 190282
Db
RESULT 9
AF141920 504 bp mRNA linear INV 31-JUL-1999
LOCUS Drosophila melanogaster roadblock (robl) mRNA, complete cds.
ACCESSION AF141920
VERSION AF141920.1 GI:5668623
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 504
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="54B"
/dev_stage="3rd instar larva"
1. 504
/gene="robl"
70, 363
/gene="robl"
/note="LC7 light chain dynein-associated protein family;
similar to Chlamydomonas outer arm axonemal dynein light
chain LC7; similar to Drosophila bithoraxoid (bxd) late
mRNA"
/codon_start=1
/product="roadblock"
/protein_id="AAD45386.1"
/db_xref="GI:5668624"
/translation="MSQEVETLKRIOHKGVGVTVNNNGIPVKSITDNTTVOYA
GLMSQADKARSVVRDLDPNSDWTFLPVRVSKKHEIIVAPDKDFILIVQNPDT"
BASE COUNT 158 a 105 c 113 g 128 t
Query Match 16.7%; Score 57.6; DB 3; Length 504;
Best Local Similarity 50.4%; Pred. No. 5e-07;
Matches 141; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
QY 41 ACGTGAAAGAGATTTTCGAAAGTGCAGAGAAACCCGCGTGGAGATATTGATCA 100
80 AGGTTGAGAAACACTCAAGAGATCCAGGCCAAGAGGTGTGTGGTACAAATGTGG 139
101 TGAATCACTCGGGTGTGCGGTGAAAACTCGATGATCGTCAAGAGGCTTCAGTACG 160
140 TCAACATGAAAGTATTTCGGTCAATTCACGCTGAGCAACACTACACCGTTCAAGTACG 199
QY 161 CTGTCTATATGACAAATTTGCGGAGAGTGCAGAGCGTTCTCTCCAAAATGAGACG 220
200 CTGGCCTAATGATGACGCTGGCGAGCAAGAGTGTGTGAGGAGCACTTGATCTT 259
QY 221 CCCAAAATTTGACTCTACTAGAGTGTGTACCAAGTATCAAGAGTCTCTATTAACACAG 280
260 CCAACGACATGACATTTCTCGGGGTGTGATCCAAAGAGCAGAGATCATGTGTGACCCG 319
QY 281 ATGCCAAGATCACCGTTTGTGTGATTCAGAAATGCCAAGA 320
320 ACAAGACTTCAATCTGATTTGATCTCAAAACCAACCGA 359
Db
RESULT 10
AY071587 541 bp mRNA linear INV 20-DEC-2001
LOCUS Drosophila melanogaster RE64145 full length cDNA.
ACCESSION AY071587
VERSION AY071587.1 GI:17946349

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
CDS
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches
141; Conservative
0; Mismatches
139; Indels
0; Gaps
0;

FLI-CDNA.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 541)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E.,
George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G.,
Miranda, A., Mungall, C.J., Munro, J., Pacled, J., Paragas, V., Park, S.,
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Rubin, G.M.
and Celniker, S.
Direct Submission
Submitted (18-DEC-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our web site
(http://fruitfly.berkeley.edu) or send email to
cna@fruitfly.berkeley.edu.
Location/Qualifiers
1..541
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/map="54C3-54C3"
1..541
/gene="trbl"
/note="alignment with genomic scaffold AB003803"
/db_xref="FLYBASE:FBgn0024196"
98..391
/gene="trbl"
/note="longest ORF"
/codon_start=1
/product="Rb64145p"
/protein_id="FAL49209.1"
/db_xref="GI:17946350"
/db_xref="FLYBASE:FBgn0024196"
/translation="MSQVEVETLKRIOSHGVGVTIYVNEGI PYKSTLNTNTVOYA
GLMSQLADKRSYVRDIPSDMTFLRVRSKRIRIMVAPDKDFILIVIQPTP"
184 a 108 c 115 g 134 t

Db 228 CTGGCCCTAATGACTGACCTGGCGACAGGACGAAAGTGTGAGGAGCTTGATCTT 287
Qy 221 CCCAAATTTGACTCTACTGAGAGTTCGTACCAAGATATCCAGGCTCATACACAG 280
Db 288 CCAACGACAGACATTCCTGCGGGTGGATCCAAAGACGAGATCATGTGTGGACCCG 347
Qy 281 ATGCAGATCACCCTGTTGGTGTTCAGAAATGCCAAGA 320
Db 348 ACAAGACTTATCTGATGTATCATCAAAACCAACCGA 387
RESULT 11
BD027650
LOCUS
DEFINITION
Sequence tag and encoded human protein.
ACCESSION
BD027650.1 GI:22569392
VERSION
JP 2001269182-A/3896.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 615)
Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
Sequence tag and encoded human protein.
Patent: JP 2001269182-A 3896 02-OCT-2001;
JOURNAL
GENSET
OS Homo sapiens (human)
PN JP 2001269182-A/3896
PD 02-OCT-2001
PR 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MIANE EDWARDS, ELMERIC DUCLAIR, JEAN YVES
FI JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers
FT CDS 158..508.
FEATURES
source
1..615
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT
172 a 165 c 154 g 116 t 8 others
ORIGIN
Query Match
Best Local Similarity
Matches
138; Conservative
0; Mismatches
142; Indels
0; Gaps
0;

15.7%; Score 54; DB 6; Length 615;
49.3%; Pred. No. 7.3e-06;
138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy 41 ACGTGAAGAAGTATTCGCAAGTGCAGAGAAACCCGCGTGAGAGACATATTCATCA 100
Db 228 AGGTGGAGGAGACACTGAAGCGACCTGCAAGNCCAGAGGAGTNCAGGAATATGCTG 287
Qy 101 TGAATCACTCGGGTGTGCGGTGAAACTCGATGATGCTCAGAGGGCTTGCACTAGC 160
Db 288 TGAACACAGAGGACATTCATCATGACAGACCATGACACACACACCACTATG 347
Qy 161 CCTGTCTATATGACATTTGCGGAGAAAGTCCAGCGCTTCTCTCAAAATGAGCCAG 220
Db 348 CCAAGCTTCATGACACCTTCATCTTGAAGAGCAAGGAGCAACGCGCTGACATGACCC 407
Qy 221 CCCAAATTTGACTCTACTGAGAGTTCGTACCAAGATATCCAGGCTCATACACAG 280
Db 408 AGACATCTCACCTTCTCTTCAATTCGCTCCAGAAATTAATTAATGATGTTGAC 467
Qy 281 ATGCCAAGATCACCGTTTGTGTGAGTTCAGAAATGCCAAGA 320
Db 468 ATAAAGACTATTTCTGATTTGATTCAGAAATCAACCGA 507

RESULT 12
AF178431 291 bp mRNA linear PRI 28-AUG-2001
LOCUS AF178431
DEFINITION Homo sapiens BITH (BITH) mRNA, complete cds.
ACCESSION AF178431
VERSION AF178431.1 GI:15320515
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 291)
TITLE Fracchiolla N.S., Corlezezi A. and Lambertenghi-Dejlliers G.
JOURNAL Bith, a human homolog of bithorax Drosophila melanogaster gene, on chromosome 20q
REFERENCE 2 (bases 1 to 291)
AUTHORS Fracchiolla N.S., Corlezezi A. and Lambertenghi-Dejlliers G.
TITLE Direct Submission
JOURNAL Submitted (18-AUG-1999) Hematology, Ospedale Maggiore Policlinico INCCS, Via Francesco Sforza 35, Milan 20122, Italy
FEATURES
source
1..291
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q; between D20S182 and D20S106"
1..291
/gene="BITH"
1..291
/note="Drosophila melanogaster bithoraxid protein homolog"
/codon_start=1
/product="BITH"
/protein_id="AAK95342.1"
/db_xref="GI:15320515"
/translation="MAVEETLKRLOSGQKVGIIIVNTGIPKSTMDPTTQYAS LMSFTLKARSTVRDIDPNDLTPFLIRSKNEIMVAPDKDYFLIVIONPTE"
BASE COUNT 91 a 81 c 67 g 52 t
ORIGIN
Query Match 15.3%; Score 52.8; DB 9; Length 291;
Best Local Similarity 49.3%; Pred. No. 1.5e-05;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 41 ACGTGGAAGAAGTATTTGCGAAGTGCAGAGAAACCGCGGTGAGGACATATTGATCA 100
DB 8 AGGTGGAGAGACACTGAAGCACTGCAGAGCCGAGAGGAGTGCAGGAAATCATCTGCG 67
QY 101 TGAATCACTCGGGTGTGCGGGTGAAGAACTTGATGATCGTCAGAGGGCTTGACATACG 160
DB 68 TGAACACAGAAAGGATTCCTCCATCAAGAGCAACATGACCAACCCACCAACAGATATG 127
QY 161 CCTGTCTATATGACAAATTTGCGGAGAGAGTCCAGGCGTCTCTCCAAATGAGCCAG 220
DB 128 CCAGCCTCATGACACAGCTTCATCTCTGAAGGACGAGAGACCGTCTGACATGACCC 187
QY 221 CCCAAATTTGACTCTACTGAGAGTTCGTACCAAGTATCAGAGGTGCTCATTAACACAG 280
DB 188 AGAAGCATCTCACTCTCTGGAATTCGCTCCAGAAAGAAATTAATGATGATGACACG 247
QY 281 ATGCCAAGTACCGGTTTGTGTGATTCAGAAATGCCAAGA 320
DB 248 ATAAAGCTATTTCTGATGTGATTCAGAAATCAACCGA 287

RESULT 13
AY026513 291 bp mRNA linear PRI 19-MAR-2001
LOCUS AY026513
DEFINITION Homo sapiens dynein-associated protein HKM23 (km23) mRNA, complete cds.
ACCESSION AY026513
VERSION AY026513.1 GI:13378303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Tang, Q., Staub, C.M. and Mulder, K.M.
TITLE km23: role in growth factor signaling
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 291)
AUTHORS Tang, Q., Staub, C.M. and Mulder, K.M.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2001) Pharmacology (H078), Penn State University College of Medicine, 500 University Dr., Hershey, PA 17033, USA
FEATURES
source
1..291
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="placenta"
1..291
/gene="km23"
1..291
/gene="km23"
/codon_start=1
/product="dynein-associated protein HKM23"
/protein_id="AAK18712.1"
/db_xref="GI:13378304"
/translation="MAVEETLKRLOSGQKVGIIIVNTGIPKSTMDPTTQYAS LMSFTLKARSTVRDIDPNDLTPFLIRSKNEIMVAPDKDYFLIVIONPTE"
BASE COUNT 91 a 81 c 67 g 52 t
ORIGIN
Query Match 15.3%; Score 52.8; DB 9; Length 291;
Best Local Similarity 49.3%; Pred. No. 1.5e-05;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 41 ACGTGGAAGAAGTATTTGCGAAGTGCAGAGAAACCGCGGTGAGGACATATTGATCA 100
DB 8 AGGTGGAGAGACACTGAAGCACTGCAGAGCCGAGAGGAGTGCAGGAAATCATCTGCG 67
QY 101 TGAATCACTCGGGTGTGCGGGTGAAGAACTTGATGATCGTCAGAGGGCTTGACATACG 160
DB 68 TGAACACAGAAAGGATTCCTCCATCAAGAGCAACATGACCAACCCACCAACAGATATG 127
QY 161 CCTGTCTATATGACAAATTTGCGGAGAGAGTCCAGGCGTCTCTCCAAATGAGCCAG 220
DB 128 CCAGCCTCATGACACAGCTTCATCTCTGAAGGACGAGAGACCGTCTGACATGACCC 187
QY 221 CCCAAATTTGACTCTACTGAGAGTTCGTACCAAGTATCAGAGGTCTCATTAACACAG 280
DB 188 AGAAGCATCTCACTCTCTGGAATTCGCTCCAGAAAGAAATTAATGATGATGACACG 247
QY 281 ATGCCAAGTACCGGTTTGTGTGATTCAGAAATGCCAAGA 320
DB 248 ATAAAGCTATTTCTGATGTGATTCAGAAATCAACCGA 287

RESULT 14
BD030271 452 bp DNA linear PAT 27-AUG-2002
LOCUS BD030271
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD030271
VERSION BD030271.1 GI:22572013
KEYWORDS JP 2001269182-A/6517.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 452)
AUTHORS Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patient: JP 2001269182-A 6517 02-OCT-2001;
GENSET

COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/6517
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDMARDS, EIMERIC DUCLAIR, JEAN YVES
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40

FEATURES
source 1.452 Location/Qualifiers.
FH Key Location/Qualifiers.
CC

BASE COUNT 129 a 124 c 107 g 91 t 1 others

ORIGIN

Query Match 15.3%; Score 52.8; DB 6; Length 452;
Best Local Similarity 49.3%; Pred. No. 1.7e-05;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 41 ACSTGAGAGAGATTTTCGCAAGTGCAGAGAGAAACCGCGCTGAGACATATTGATCA 100
DB 65 AGGTGAGAGAGACATGAAAGCGACTGCAAGCCAGAGAGATGCGAGATCATGCTG 124
QY 101 TGAATCACTCGGGGTGCGCGGTGAAAACCTCGATGATGCTGAGAGGGCTTCAGTAGC 160
DB 125 TGAACACAGAGAGCATTTCCCATCAAGAGCAGCATGAGCAACCCACACACCAAGTAGT 184
QY 161 CCTGTATATGACATTTTGGGGAGAGAGTCCAGAGGCTTCTCCAAATGAGCCAG 220
DB 185 CAGCCTCATGACAGCTTCATCTGAAAGCAGAGACAGCGCTGACATGACCCCC 244
QY 221 CCCAAATTTGACTCTACTGAGAGTTCGTAAGATTCAGAGGTCTCATTAACAG 280
DB 245 AGAAGCATCTCACTTCTTCGATTCGCTCCAAAGAAAATGAATTAATGTTGACACAG 304
QY 281 ATGCCAAGATCACCGTTTGGTTCAGATGCCAAGA 320
DB 305 ATTAAGACTATTTCTGATTTGATTCAGATCCAAACGA 344

RESULT 15
BD111742 603 bp DNA linear PAT 18-SEP-2002
LOCUS EST and encoded human protein.
DEFINITION BD111742
ACCESSION BD111742.1 GI:23206560
VERSION JP 2002010789-A/3819.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Bukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 603)
AUTHORS Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.
TITLE EST and encoded human protein
JOURNAL Patient: JP 2002010789-A 3819 15-JAN-2002;
GENSET CORP

COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/3819
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDMARDS, SEVELIN JOBERT, JEAN EVE PI

GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT CDS 68..355.
Location/Qualifiers
source 1.603
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 156 a 182 c 146 g 119 t

ORIGIN

Query Match 15.3%; Score 52.8; DB 6; Length 603;
Best Local Similarity 49.3%; Pred. No. 1.8e-05;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 41 ACSTGAGAGAGATTTTCGCAAGTGCAGAGAGAAACCGCGCTGAGACATATTGATCA 100
DB 75 AGGTGAGAGAGACATGAAAGCGACTGCAAGCCAGAGAGATGCGAGATCATGCTG 134
QY 101 TGAATCACTCGGGGTGCGCGGTGAAAACCTCGATGATGCTGAGAGGGCTTCAGTAGC 160
DB 135 TGAACACAGAGAGCATTTCCCATCAAGAGCAGCATGAGCAACCCACACCAAGTAGT 194
QY 161 CCTGTATATGACATTTTGGGGAGAGTCCAGAGGCTTCTCCAAATGAGCCAG 220
DB 195 CAGCCTCATGACAGCTTCATCTGAAAGCAGAGACAGCGCTGACATGACCCCC 254
QY 221 CCCAAATTTGACTCTACTGAGAGTTCGTAAGATTCAGAGGTCTCATTAACAG 280
DB 255 AGAAGCATCTCACTTCTTCGATTCGCTCCAAAGAAAATGAATTAATGTTGACACAG 314
QY 281 ATGCCAAGATCACCGTTTGGTTCAGATGCCAAGA 320
DB 315 ATTAAGACTATTTCTGATTTGATTCAGATCCAAACGA 354

Search completed: February 8, 2004, 04:40:20
Job time : 1321.21 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 7, 2004, 18:33:25 ; Search time 98.295 Seconds
(without alignments)
9474.605 Million cell updates/sec

Title: US-09-614-150A-2

Perfect score: 345
Sequence: 1 atcgaggaagcagagacaga.....ttcttaacataaagggttag 345

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: N.Geneseq.19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	345	100.0	345	23	ABL01841	Drosophila melanog
2	325.2	94.3	2397	23	ABL01840	Drosophila melanog
3	80.6	23.4	2635	23	ABL16802	Drosophila melanog
4	57.6	16.7	504	23	ABL16803	Drosophila melanog
5	54	15.7	615	21	AA003905	Human secreted pro
6	52.8	15.3	452	21	AA006526	Human secreted pro
7	52.8	15.3	665	22	AAH78734	Human HSP159 seri
8	52.8	15.3	682	24	AB060899	HSP162 encoding s

C	9	52.8	15.3	740	20	AA241232	Human normal ovari
	10	52.8	15.3	858	21	AAF18164	Lung cancer associ
	11	52.4	15.2	396	22	AAF94861	Human ovarian canci
	12	52.4	15.2	396	24	ABT03128	Human ovarian carc
	13	52.4	15.2	396	24	ABL48811	Ovarian carcinoma
	14	51.8	15.0	650	24	AB060352	Human colon cancer
	15	51.2	14.8	763	22	AAH26540	Human mediator of
	16	50.8	14.7	294	23	ABL16875	Drosophila melanog
	17	50.8	14.7	2294	23	ABL16874	Drosophila melanog
	18	49	14.2	2635	23	ABL16802	Drosophila melanog
	19	49	14.2	9142	23	ABL20376	Drosophila melanog
	20	49	14.2	9142	23	ABL25232	Drosophila melanog
	21	46.4	13.4	749	22	AAH26541	Rat mediator of TG
	22	43.2	12.5	294	23	ABL16871	Drosophila melanog
	23	43.2	12.5	2294	23	ABL16870	Drosophila melanog
	24	42.2	12.2	3539	23	ABL16814	Drosophila melanog
	25	40.6	11.8	574	22	ABA63884	Human foetal liver
	26	40.6	11.8	574	22	ABA30874	Probe #9340 for ge
	27	40.6	11.8	574	22	AAK12200	Human brain expres
	28	40.6	11.8	574	22	AAK37922	Human bone marrow
	29	40.6	11.8	574	22	AA116883	Probe #12482 used
	30	40.6	11.8	574	22	AA143796	Human liver single
	31	40.6	11.8	574	23	ABS37544	Human genome-deriv
	32	40.6	11.8	574	24	ABS11918	Drosophila melanog
	33	38.8	11.2	2464	23	ABL18474	Drosophila melanog
	34	38.8	11.2	2903	23	ABU06252	Drosophila melanog
	35	38	11.0	393	23	ABL18475	Drosophila melanog
	36	37.4	10.8	3176	21	AACT6094	Human ORF1643
	37	36.4	10.6	619	23	ABK41859	CDNA encoding nove
	38	35	10.1	1437	23	ABL15419	Drosophila melanog
	39	35	10.1	3437	23	ABL15418	Drosophila melanog
	40	33.4	9.7	1316	21	AACT75192	Human ORF16747
	41	33.4	9.6	3156	25	AA053174	Human kinases and
	42	33	9.6	2528	23	ABD09043	Drosophila melanog
	43	33	9.6	5309	23	ABL09023	Drosophila melanog
	44	33	9.6	7439	23	ABL09042	Drosophila melanog
	45	33	9.6	20629	23	ABL09022	Drosophila melanog

ALIGNMENTS

RESULT 1	ABL01841	standard; cDNA; 345 BP.
ID	ABL01841	
AC	ABL01841;	
XX		
DT	26-MAR-2002 (first entry)	
XX		
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 5.	
XX		
KW	Drosophila; developmental biology; cell signalling; insecticide;	
KW	pharmaceutical; gene; ss.	
XX		
OS	Drosophila melanogaster.	
XX		
PN	WO200171042-A2.	
XX		
PD	27-SEP-2001.	
XX		
PP	23-MAR-2001; 2001WO-US09231.	
XX		
PR	23-MAR-2000; 2000US-191637P.	
XX		
PR	11-JUL-2000; 2000US-0614150.	
XX		
PA	(PEKE) PE CORP NY.	
XX		
PI	Venter JC, Adams M, Li PMD, Myers EW;	
XX		
DR	WPI; 2001-656860/75.	
XX		
DR	P-PSDB; ABB57738.	
XX		

PD 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2000; 2000US-191637P.
 XX 11-JUL-2000; 2000US-0614150.
 PR (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW.
 PI WPI; 2001-656860/75.
 DR
 XX
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX
 PS Claim 1; SEQ ID NO 1879; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB16175), expressed DNA
 CC sequences (AB161840-AB16175) and the encoded proteins
 CC (AB16173-AB16172).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 2635 BP; 793 A; 508 C; 494 G; 840 T; 0 other;
 Query Match 23.4%; Score 80.6; DB 23; Length 2635;
 Best Local Similarity 53.7%; Pred. No. 6.7e-17;
 Matches 167; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
 QY 16 ACAGATCAAGAGCAATTAAGTATGCTGAGAGGATTTGCGCAAGTGCAGAGAA 75
 DB 2053 AATATGCACTCTGCACTTACATATGTGAAAAACATTCGACTGTCATATCGAAA 2112
 QY 76 CCGCGCGTGGAGGACATATTTGATCATGATCACTCGGATGTCGGTGAAGCCTCGATG 135
 DB 2113 AATCATTTGAGATGATGATGATCTCAAGAGTCCGGAGATCCGGATGATGATGATG 2172
 QY 136 GATGCTAGAGAGGCTTGCAGTACGCTGTCTATATGACAATTTGCGGAGAAAGTCCAG 195
 DB 2173 GACCGAGAGGATGCTGTCATTTTCGGGACCTTTTCAGGCTATTCAGAGAAAGTGGAA 2232
 QY 196 GCGTTCCTCTCCAAATGAGGAGCCGCAAAATTTGACTCTACGTAGAGAGTTCGACCAAG 255
 DB 2233 CGCGGATGATGATTAATGATTCACGAGAGAACTAAATGCTGAGAAATGACGACAGA 2292
 QY 256 TATACAGAGGTGCTCATTAACAGATCCAAAGATCACTTTTGTGTGTTCAAGATGCC 315
 DB 2293 ACCAAGAGGTGCTCCCTGCGGACGAGTCCAAAGTAACTACTCTGATGATGAGAGCGC 2352
 QY 316 AAAGATTAATTT 326
 DB 2353 CATGATTACTT 2363
 RESULT 4
 ID ABL16803
 ID ABL16803 standard; DNA; 504 BP.
 AC ABL16803;
 XX
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 1882.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.
 OS
 XX Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW.
 PI WPI; 2001-656860/75.
 DR
 XX
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX
 PS Claim 1; SEQ ID NO 1882; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB16175), expressed DNA
 CC sequences (AB161840-AB16175) and the encoded proteins
 CC (AB16173-AB16172).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 504 BP; 160 A; 103 C; 113 G; 128 T; 0 other;
 Query Match 16.7%; Score 57.6; DB 23; Length 504;
 Best Local Similarity 50.4%; Pred. No. 2.5e-09;
 Matches 141; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
 QY 41 AGGTGAGAGAGATTTCCGAAGTGCAGAGAAACCGGCGTGGAGATTTGATCA 100
 DB 80 AGGTGAGAGAGATTTCCGAAGTGCAGAGAAACCGGCGTGGAGATTTGATCA 139
 QY 101 TGAATCACTCGGATGTCGGGAGAAACCTCGATGATGTCAGAGGCTTGCAGTACG 160
 DB 140 TCAACATGAGAGTATTTCCGGTCAATTCACGCTGAGCAACACTACACCGTTCACTGACG 199
 QY 161 CCGTCTATATGACAAATTTGCGGAGAGATGTCAGGCGTTCTCTCCAAATGAGACCG 220
 DB 200 CTGGCTTAATGAGTCACTGCGGAGCAAGGACGAAATGATGAGAGGACTTGAATCCTT 259
 QY 221 CCCAAATTTGACTCTACGAGAGTTCGATCAAGATGATGAGAGGCTCATTAACCGAG 280
 DB 260 CCAAGCAATGAGATTCCTGCGGAGTCAATCCAAAGAGCAAGATCAATGATGAGACCGC 319
 QY 281 ATGCCAAGATCAACCGTTTGTGTGTTGATGAGATGAGATGAGATGAGATGAGATGAG 320
 DB 320 ACAAGGACTTCACTCTGATTTGATTCAAACCAACCAACGCA 359
 RESULT 5
 ID AAC03905
 ID AAC03905 standard; cDNA; 615 BP.
 AC AAC03905;
 XX
 XX 06-OCT-2000 (first entry)
 DE Human secreted protein 5' EST, SEQ ID NO: 3903.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX P-PSDB; AAG03899.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 3903; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
CC
XX
SQ Sequence 615 BP; 172 A; 165 C; 154 G; 116 T; 8 other;

Query Match 15.7%; Score 54; DB 21; Length 615;
Best Local Similarity 49.3%; Pred. No. 4.9e-08;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 41 ACGTGAAGAAGTATTTGCGAAGTGCAGAGAAACCGGCGTGGAGGACATATTGATCA 100
DB 228 AGGTGGAGAGACACTGAAGGACTGCGANCCAGAGAGAGTMCAGGATCATCTGCG 287
QY 101 TGAATCACTGGGTGTCGGGTGAAAACCTTGATGATGTCAGAGGGCTTGCAGTACG 160
DB 288 TGAACACGAAGGATTCCTCCATCAAGACACCATGACCAACACACCCAGATG 347
QY 161 CCTGTATATGACAAATTTGGGAGAGAGTCCAGGCGTCTCTCCAAATGAGACG 220
DB 348 CCACCTCATGACACAGTTTCATCTGAAAGGACGAGAGACCGTGCATGACACCCCC 407
QY 221 CCCAAATTTGACTCTACTGAGAGTTCGTAACCAAGTATCAAGAGTGTCTATTACACG 280
DB 408 AGAAGCATCTCACTTCTTGAATTCGCTCCAGAAAATGAAATTTGTTGACACG 467
QY 281 ATGCCAAGATCACCGTTTGTGTTGTTCAAGATGCCAAGA 320
DB 468 ATAAAGACTATTTCTGATGTGATTCAGATTCACGCA 507

RESULT 6
AAC06526
ID AAC06526 standard; cDNA; 452 BP.
XX

AC AAC06526;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 10601.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 10601; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
CC
XX
SQ Sequence 452 BP; 129 A; 124 C; 107 G; 91 T; 1 other;

Query Match 15.3%; Score 52.8; DB 21; Length 452;
Best Local Similarity 49.3%; Pred. No. 1.1e-07;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 41 ACGTGAAGAAGTATTTGCGAAGTGCAGAGAAACCGGCGTGGAGGACATATTGATCA 100
DB 65 AGGTGGAGAGACACTGAAGGACTGCGAGGCCAGAGAGAGTGCAGGATCATCTGCG 124
QY 101 TGAATCACTGGGTGTCGGGTGAAAACCTTGATGATGTCAGAGGGCTTGCAGTACG 160
DB 125 TGAACACGAAGGATTCCTCCATCAAGACACCATGACCAACACCCAGATG 184
QY 161 CCTGTATATGACAAATTTGGGAGAGAGTCCAGGCGTCTCTCCAAATGAGACG 220
DB 185 CCAGCTCATGACACAGTTTCATCTGAAAGGACGAGAGACCGTGCATGACACCCCC 244
QY 221 CCCAAATTTGACTCTACTGAGAGTTCGTAACCAAGTATCAAGAGTGTCTATTACACG 280
DB 245 AGAAGCATCTCACTTCTTGAATTCGCTCCAGAAAATGAAATTTGTTGACACG 304
QY 281 ATGCCAAGATCACCGTTTGTGTTGTTCAAGATGCCAAGA 320
DB 305 ATAAAGACTATTTCTGATGTGATTCAGATTCACGCA 344

RESULT 7
AAH78734 standard; cDNA; 665 BP.
XX
XX AAH78734;
XX
XX 03-DEC-2001 (first entry).
XX
XX Human HsIFL59 serine/threonine phosphatase cDNA sequence.
XX
XX Human; HsIFL59; ss; serine/threonine phosphatase; PSPase; vaccine;
XX gene therapy; PSPase expression; PSPase expression; PSPase modulation;
XX immune disorder; autoimmune disorder; Wiscott-Aldrich syndrome;
XX Chediak-Higashi syndrome; Hashimoto's thyroiditis; multiple sclerosis;
XX inflammation; Crohn's disease; inflammatory bowel disease; appendicitis;
XX rheumatoid arthritis; cellular proliferative disorder; lymphoma;
XX lung cancer; intestinal cancer; cardiovascular disorder; aneurysm;
XX Scimitar syndrome; Ebstein's anomaly.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 24..314
XX /*tag= a
XX /product= "HsIFL59 PSPase protein"
XX
XX W0200164703-A1.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US06256.
XX
XX 02-MAR-2000; 2000US-0186350.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ebner R, Ruben SM;
XX
XX WPI; 2001-530113/58.
XX
XX P-PSDB; AAG77807.
XX
XX Nucleic acids encoding serine/threonine phosphatase polypeptides,
XX useful for preventing, diagnosing and/or treating, e.g. Crohn's
XX disease, lung cancer and Scimitar syndrome -
XX
XX Claim 1; Page 310; 335pp; English.
XX
XX The present sequence represents the specifically claimed human cDNA clone
XX HsIFL59 which shares homology with members of the serine/threonine family
XX of phosphatases. The invention comprises novel human serine/threonine
XX phosphatase (PSPase) polypeptides and polynucleotides which may be used
XX in the prevention (vaccine), diagnosis and treatment (gene therapy) of
XX diseases associated with inappropriate PSPase expression. The PSPase
XX polynucleotides of the invention may be used as DNA probes to detect and
XX quantitate the presence of similar nucleic acids in samples. The PSPase
XX polypeptides may be used as antigens in the production of antibodies
XX against the PSPase polypeptides and in assays to identify modulators of
XX PSPase expression and activity. The anti-PSPase antibodies and
XX antagonists may also be used to down regulate expression and activity,
XX the anti-PSPase antibodies may also be used as diagnostic agents for
XX detecting the presence of PSPase polypeptides in samples. Disorders that
XX may be prevented, diagnosed and/or treated by the invention are:
XX immune/autoimmune disorders (e.g. Wiscott-Aldrich syndrome,
XX Chediak-Higashi syndrome, Hashimoto's thyroiditis and multiple
XX sclerosis); inflammatory conditions (e.g. Crohn's disease, inflammatory
XX bowel disease, appendicitis and rheumatoid arthritis); cellular
XX proliferative disorders (e.g. lymphoma, lung and intestinal cancers); and
XX cardiovascular disorders (e.g. Scimitar syndrome, Ebstein's anomaly and
XX aneurysm).
XX
XX Sequence 665 BP; 187 A; 195 C; 143 G; 140 T; 0 other;
XX
XX Query Match 15.3%; Score 52.8; DB 22; Length 665;

Best Local Similarity 49.3%; Pred. No. 1.3e-07;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 41 ACCTGGAAGAAGATTTTCCGCAAGTGCAGAGAAACCCGCGTGGAGGACATTTGATCA 100
DB 31 AGGTGAGGAGACACGAAACCGACGACAGCCAGAGAGGAGGAAATCATGTCG 90
QY 101 TGAATCATCTGGGTGCGCGGTGAAGAACTGATGATGTCAGAGAGGGCTTCAGTACG 160
DB 91 TGAACACAGAGGCAATTTCCATCAAGAGACCATGACCAACCCACACCACTGATG 150
QY 161 CTTGTATATATGCAATTTGCGGAGAGTCCAGGCGTTCTCTCCAAATGAGCCAG 220
DB 151 CCAGCTCATGACACCTTCATCTGAAAGGACGAGACCGGTGCGATGACGACCC 210
QY 221 CCCAAATTTGACTCTACGAGAGTTCGACCAAGATTCACGAGGTCTCATTCACCG 280
DB 211 AGAACATCTCACCTTCCTTCGAATTCGCTCCAAAGAAATTAATGTTGACACAG 270
QY 281 ATGCCAAGATCACCGTTTGTGTTGTCAGAAATGCCAAAG 320
DB 271 ATAAAGACTATTTCTGATTTGATTCAGATTCAAACCA 310
RESULT 8
ABQ60899 standard; cDNA; 682 BP.
XX
XX ABQ60899;
XX
XX 26-FEB-2003 (first entry)
XX
XX HSPCL62 encoding sequence.
XX
XX Neuroprotective; immunomodulator; cancer; chromosome 20pter-p12.3;
XX cytostatic; anti-inflammatory; gene therapy; nutritional supplement;
XX wound; burn; ulcer; Alzheimer's disease; Huntington's disease;
XX amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
XX vulnerary; gene; ss.
XX
XX Homo sapiens.
XX
XX W0200231111-A2.
XX
XX 18-APR-2002.
XX
XX 11-OCT-2001; 2001WO-US27760.
XX
XX 12-OCT-2000; 2000US-0687527.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX
XX Tang AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-426278/45.
XX
XX N-PSDB; ABP43655.
XX
XX New polypeptides and their encoded proteins, useful as nutritional
XX sources or supplements, or in gene therapy, particularly for treating
XX wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
XX inflammation -
XX
XX Claim 1; SEQ ID # 112; 357pp + sequence listing; English.
XX
XX The invention relates to 446 newly isolated polynucleotide sequences.
XX The activity of polynucleotides of the invention may be described as,
XX vulnerary, neuroprotective, immunomodulator, cyostatic and
XX anti-inflammatory. Compositions comprising nucleic acids of the invention
XX are useful for treating a mammalian subject, or as nutritional sources or
XX supplements. These are useful in gene therapy, particularly for treating
XX wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
XX amyotrophic lateral sclerosis, autoimmune disorders, cancer or

CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records
CC AB060788-AB061233 represent polynucleotides of the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 682 BP; 186 A; 200 C; 149 G; 147 T; 0 other;

Query Match 15.3%; Score 52.8; DB 24; Length 682;
Best Local Similarity 49.3%; Pred. No. 1.3e-07;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 41 ACGTGAAGAAGTATTCGCAAGTGCAGAGAAACCCGGCGTGCAGACATATTGATCA 100
DB 55 AGGTGAGAGAGACACTGAAGCACTGCAGAGCCGAAAGGAGTGCAGGAATCATGCTCG 114
QY 101 TGAATCACTCGGGTGTGCGGGTGAACCTCGATGATCGTCAGAGGGCTTGCAATACG 160
DB 115 TGAACACAGAGAGCATTCCTCATGAAGACACATGACCAACCCACACCCAGATATG 174
QY 161 CCGTCTATATGACAAATTTGGCGGAGAAAGTCCAGGGGTTCTCTCAAAATGAGCCAG 220
DB 175 CCAAGCTCATGACACAGCTTCACTCTGAAGGACGAGACACCGTGCATGACATGACCCC 234
QY 221 CCCAAATTTGACTCTACTGAGAGTGTGACCAAGTATCAGAGGTGCTCATTAACACAG 280
DB 235 AGAAGCATCTCACCTTCTCTGCAATTCGCTTCAAGAAATGAAATTAATGTTGCACACG 294
QY 281 ATGCCAAGATCACCGTTTGGTGGTTCAGATGCCAAGA 320
DB 295 ATAAAGACTATTTCTGATTTGTGATTCAGATCAACCGA 334

RESULT 9

AAZ41232/c
ID AAZ41232 standard; cDNA; 740 BP.

XX
AC AAZ41232;

XX
DT 18-JAN-2000 (first entry)

XX
DE Human normal ovarian tissue derived cDNA 11.

XX
KW Human; ovary; screening; ovarian cancer; treatment; ss.

OS Homo sapiens.

XX
PN DE19816395-A1.

XX
PD 07-OCT-1999.

XX
PF 03-APR-1998; 98DE-1016395.

XX
PR 03-APR-1998; 98DE-1016395.

XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX
DR WPI; 1999-552352/47.

XX
PT P-PSDB; AAY59743, AAY59744, AAY59745.

XX
PT Nucleic acid sequences potentially useful in diagnosis or therapy of
XX ovarian cancer

XX
PS Claim 3; Page 129; 274pp; German.

XX
CC This invention describes novel nucleic acid sequences that are highly
XX expressed in normal ovary tissue. Artificial chromosomes and cosmid
XX clones containing the sequences can be used as gene transfer vehicles.
XX The sequences can be used to produce DNA fragments containing
XX full-length genes. Host cells transformed with the sequences can be used

CC to produce polypeptides or polypeptide fragments, which can be used to
CC screen phage displays for polypeptides that bind to them, or as tools for
CC identifying agents active against ovarian cancer, or to prepare
CC medicaments for treating ovarian cancer. The cDNA sequences can be used
CC to obtain genomic genes, their promoters, enhancers, silencers, exon
CC structures, intron structures and their splice variants. AAZ41222-241324
CC represent cDNA sequences derived from normal human ovarian tissue and
CC which encode the protein fragments represented in AAY59724-Y59837.

XX
SQ Sequence 740 BP; 150 A; 180 C; 205 G; 205 T; 0 other;

Query Match 15.3%; Score 52.8; DB 20; Length 740;
Best Local Similarity 49.3%; Pred. No. 1.4e-07;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 41 ACGTGAAGAAGTATTCGCAAGTGCAGAGAAACCCGGCGTGCAGACATATTGATCA 100
DB 673 AGGTGAGAGAGACACTGAAGCACTGCAGAGCCGAAAGGAGTGCAGGAATCATGCTCG 614
QY 101 TGAATCACTCGGGTGTGCGGGTGAACCTCGATGATCGTCAGAGGGCTTGCAATACG 160
DB 613 TGAACACAGAGAGCATTCCTCATGAAGACACATGACCAACCCACACCCAGATATG 554
QY 161 CCGTCTATATGACAAATTTGGCGGAGAAAGTCCAGGGGTTCTCTCAAAATGAGCCAG 220
DB 553 CCAAGCTCATGACACAGCTTCACTCTGAAGGACGAGACACCGTGCATGACATGACCCC 494
QY 221 CCCAAATTTGACTCTACTGAGAGTGTGACCAAGTATCAGAGGTGCTCATTAACACAG 280
DB 493 AGAAGCATCTCACCTTCTCTGCAATTCGCTTCAAGAAATGAAATTAATGTTGCACACG 434
QY 281 ATGCCAAGATCACCGTTTGGTGGTTCAGATGCCAAGA 320
DB 433 ATAAAGACTATTTCTGATTTGTGATTCAGATCAACCGA 394

RESULT 10

AAF18164
ID AAF18164 standard; DNA; 858 BP.

XX
AC AAF18164;

XX
DT 14-MAR-2001 (first entry)

XX
DE Lung cancer associated polynucleotide sequence SEQ ID 183.

XX
KW Human; lung cancer associated protein; neuroprotective; cytoprotective;
XX cardioactive; immunomodulatory; muscular active; vulnereary;

XX
KW gastrointestinal; nephrotoxic; antiinfective; gynecological;
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX proliferative disorder; wound healing; infectious disease; ds.

XX
OS Homo sapiens.

XX
PN W0200055180-A2.

XX
PD 21-SEP-2000.

XX
PF 08-MAR-2000; 2000MO-US05918.

XX
PR 12-MAR-1999; 99US-0124270.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PA (ROSE/) ROSEN C A.

XX
PI Ruben SM;

XX
DR WPI; 2000-587514/55.

XX
PT P-PSDB; AAB58288.

XX
CC Lung cancer associated gene sequences, referred to as lung cancer
XX antigens, useful for treatment, prevention, and diagnosis of disorders
XX such as lung cancer

XX Claim 1; Page 647; 1425bp; English.
XX
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnary; gastrointestinal
CC general; nephrotoxic; antifibrotic; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterization of the polynucleotide and protein
CC sequences.
XX
XX Sequence 858 BP; 221 A; 244 C; 206 G; 184 T; 3 other;
SQ
Query Match 15.3%; Score 52.8; DB 21; Length 858;
Best Local Similarity 49.3%; Pred. No. 1.5e-07;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 41 ACGTGAAGAAGTATTTCCGAAAGTGCAGAGAAACCCGCGTGAGAGACATATTGATCA 100
DB 197 AGGTGGAGAGACACTGAAACCGACTGCAGAGCCGAGAGAGTGCAGAGATCATCTGTCG 256
QY 101 TGAATCACTCGGGTGTGCCGGAACCTCGATGATGCTCAGAGAGGCTTGAGATAGC 160
DB 257 TGAACACAGAGGATTCATCAAGACACCAATGACAAACCCACACACCCAGATG 316
QY 161 CCTGTCTATATGACAAATTTGCGGAGAGTCCAGGCGTCTCTCCAAATGAGCCAG 220
DB 317 CCAGCCTCATGACAGCTTCTCATCTGAAAGCAGGACACCGTCCGATGATGACCC 376
QY 221 CCCAAATTTGACTCTACTGAGAGTTCGTAACAAATATCAAGAGTCTCATTAACCCAG 280
DB 377 AGAACATCTCACCTTCTTCAATTCGCTCAAGAAATGAATATGATGTCACACG 436
QY 281 ATGCCAAGATCACCGTTTGTGTGTTGATGATGCAATGCCAAGA 320
DB 437 ATAAAGACTATTTCTGATTTGATTCAGATCCAAACCGA 476

RESULT 11
AAF94861
ID AAF94861 standard; cDNA; 396 BP.
XX
XX AAF94861;
AC
XX
XX 23-MAY-2001 (first entry)
DT
XX
XX Human ovarian cancer associated coding sequence SEQ ID NO: 52.
DE
XX
XX Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200118046-A2.
PN
XX
XX 15-MAR-2001.
PD
XX
XX 08-SEP-2000; 2000WO-US24827.
PF
XX
XX 10-SEP-1999; 99US-0394374.
PR
XX
XX 01-MAY-2000; 2000US-0561778.
PR
XX
XX 15-AUG-2000; 2000US-0640173.
PR
XX
XX 07-SEP-2000; 2000US-0656668.

XX (CORI-) CORIXA CORP.
PA
XX
XX Xu J, Stolk JA;
PI
XX
XX WPI; 2001-211395/21.
DR
XX
XX Isolated polypeptides associated with ovarian carcinomas, and the
PT nucleic acids that encode them, useful for the prevention diagnosis and
PT treatment of ovarian cancers -
PS
XX
XX Claim 5; Page 132; 1899p; English.
XX
XX The present invention provides a number of coding sequences and proteins,
CC the over-expression of which is associated with ovarian carcinoma/cancer.
CC These can be used in the diagnosis, treatment and prevention of ovarian
CC cancer, optionally by gene therapy or in the form of a vaccine. The
CC present sequence is an example of one of these sequences.
XX
XX Sequence 396 BP; 111 A; 111 C; 90 G; 82 T; 2 other;
SQ
Query Match 15.2%; Score 52.4; DB 22; Length 396;
Best Local Similarity 48.9%; Pred. No. 1.4e-07;
Matches 137; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 41 ACGTGAAGAAGTATTTCCGAAAGTGCAGAGAAACCCGCGTGAGAGACATATTGATCA 100
DB 53 AGGTGAAGAGACACTGAAACCGACTGCANAGCCAGAAAGAGTGCAGAGATCATCTGTCG 112
QY 101 TGAATCACTCGGGTGTGCCGGAACCTCGATGATGCTCAGAGAGGCTTGACATAGC 160
DB 113 TGAACACAGAGGATTCATCCATCAAGACACCATGACAAACCCACACCCAGATG 172
QY 161 CCTGTCTATATGACAAATTTGCGGAGAGTCCAGGCGTCTCTCCAAATGAGCCAG 220
DB 173 CCAGCCTCATGACAGCTTCTCATCTGAAAGCAGGACACCGTCCGATGATGACCC 232
QY 221 CCCAAATTTGACTCTACTGAGAGTTCGTAACAAATATCAAGAGTCTCATTAACCCAG 280
DB 233 AGAACATCTCACCTTCTTCAATTCGCTCAAGAAATGAATATGATGTCACACG 292
QY 281 ATGCCAAGATCACCGTTTGTGTGTTGATGATGCAATGCCAAGA 320
DB 293 ATAAAGACTATTTCTGATTTGATTCAGATCCAAACCGA 332

RESULT 12
ABT03128
ID ABT03128 standard; cDNA; 396 BP.
XX
XX ABT03128;
AC
XX
XX 05-SEP-2002 (first entry)
DT
XX
XX Human ovarian carcinoma associated coding sequence SEQ ID NO: 52.
DE
XX
XX Human, ovarian cancer; ovarian carcinoma; gene therapy; immunotherapy;
KW cytostatic; gene; ss.
XX
XX
XX Homo sapiens.
OS
XX
XX WO200239885-A2.
PN
XX
XX 23-MAY-2002.
PD
XX
XX 13-NOV-2001; 2001WO-US45395.
PF
XX
XX 14-NOV-2000; 2000US-0713550.
PR
XX
XX 03-APR-2001; 2001US-0825294.
PR
XX
XX 02-OCT-2001; 2001US-0970966.
XX
XX (CORI-) CORIXA CORP.
PA
XX

PI Xu J, Stolk JA, Algate PA, Fling SP, Molesh DA;
XX WPI; 2002-500166/53.
XX
PT Novel ovarian cancer polypeptide and polynucleotide, useful for
XX detecting the presence of ovarian cancer in a patient, and in
PT pharmaceutical compositions, e.g. vaccines, for treating ovarian cancer
XX
XX Example 1; Page 130; 197pp; English.
XX
XX The present invention provides human ovarian cancer associated proteins
CC and coding sequences. The sequences can be used in the diagnosis and
CC treatment of ovarian cancers. The present sequence is a coding sequence
CC of the invention.
XX
SQ Sequence 396 BP; 111 A; 111 C; 90 G; 82 T; 2 other;
Query Match 15.2%; Score 52.4; DB 24; Length 396;
Best Local Similarity 48.9%; Pred. No. 1.4e-07;
Matches 137; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 41 ACCTGGAAGAATTTTCGCAAGTGCAGAGAAACCCGCGTGGAGCATATTGATCA 100
DB 53 AGGTGGAGAGACACTGAAAGCACTGCANAGCCGAAAGGAGTGCAGGATATCGTCG 112
QY 101 TGAATCACTGGGGTGGCCGGTGAAGAACTCGATGATCGTACAGAGGGCTTGACATACG 160
DB 113 TGAACACAGAAAGGCAATTCCTCATCAAGACACCATGACCAACCCACCACTATG 172
QY 161 CCTGTCTATATGACAAATTTGGGGAGAGTCCAGGCGTCTCTCCAAATATGAGCCAG 220
DB 173 CCAGCCTCATGACAGATTTATCTCTGAAGGACGAGACACCGTGCATGACATGACCCC 232
QY 221 CCCAAATTTGACTCTACTGAGAGTTGTAACCAAGTATCAGAGGTGCTCATTCACACG 280
DB 233 AGAAGCATCTCACCTTCCTTCGAAATTCGCTCCAAAGAAATTAATGATGTCACACG 292
QY 281 ATGCCAAGATCACCGTTTGGTGGTGCAGATGCCAAGA 320
DB 293 ATAAAGACTATTTCTGATTTGTGATTCAGATCCACCGA 332

RESULT 13
ABL48811
ID ABL48811 standard; cDNA; 396 BP.
XX
AC ABL48811;
XX
DT 18-JUN-2002 (first entry)
XX
DE Ovarian carcinoma sequence isolate 24367.
XX
KM Ovarian cancer; cancer therapy; vaccine; gene therapy; tumour; cancer;
XX ss.
XX Homo sapiens.
XX OS
XX US200204491-A1.
XX
PD 10-JAN-2002.
XX
PF 03-APR-2001; 2001US-0825294.
XX
PR 10-SEP-1999; 99US-0394374.
XX
PR 01-MAY-2000; 2000US-0561778.
XX
PR 15-AUG-2000; 2000US-0640173.
XX
PR 07-SEP-2000; 2000US-0656668.
XX
PR 14-NOV-2000; 2000US-0713550.
XX
XX (XUJ/J) XU J.
XX PA (STOL/) STOLK J A.
XX PA (ALGA/) ALGATE P A.

PA (FLIN/) FLING S P.
PI Xu J, Stolk JA, Algate PA, Fling SP;
XX WPI; 2002-171027/22.
XX
DR Ovarian tumour polypeptide and polynucleotide useful in diagnosis,
XX prevention and/or treatment of cancer, especially ovarian cancer
PT
XX Claim 1a; Page 56; 131pp; English.
XX
XX The invention relates to ovarian tumour polynucleotides and polypeptides
CC that may be utilised in cancer therapy, for example in a vaccine or
CC gene therapy. Polypeptides and polynucleotides of the invention are
CC useful for detecting a cancer in a patient, for stimulating and/or
CC expanding T-cells specific for a tumour protein, and for inhibiting the
CC development of a cancer in a patient. They are also useful for
CC stimulating an immune response in a patient, and for treating a cancer in
CC a patient and for determining the presence of a cancer in a patient.
CC The isolated polynucleotides of the invention are useful for their
CC ability to selectively form duplex molecules with complementary stretches
CC of the entire desired gene or gene fragments, and for designing and
CC preparing ribozyme molecules for inhibiting expression of tumour
CC polypeptides in tumour cells. Polypeptides and polynucleotides of the
CC invention are also useful in recombinant DNA molecules to direct
CC expression of a polypeptide in appropriate host cells. The sequences
CC given in records ABL48760-ABL48956 represent polynucleotides encoding
XX ovarian carcinoma proteins.
SQ Sequence 396 BP; 111 A; 111 C; 90 G; 82 T; 2 other;
Query Match 15.2%; Score 52.4; DB 24; Length 396;
Best Local Similarity 48.9%; Pred. No. 1.4e-07;
Matches 137; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 41 ACCTGGAAGAATTTTCGCAAGTGCAGAGAAACCCGCGTGGAGCATATTGATCA 100
DB 53 AGGTGGAGAGACACTGAAAGCACTGCANAGCCGAAAGGAGTGCAGGATATCGTCG 112
QY 101 TGAATCACTGGGGTGGCCGGTGAAGAACTCGATGATCGTACAGAGGGCTTGACATACG 160
DB 113 TGAACACAGAAAGGCAATTCCTCATCAAGACACCATGACCAACCCACCACTATG 172
QY 161 CCTGTCTATATGACAAATTTGGGGAGAGTCCAGGCGTCTCTCCAAATATGAGCCAG 220
DB 173 CCAGCCTCATGACAGATTTATCTCTGAAGGACGAGACACCGTGCATGACATGACCCC 232
QY 221 CCCAAATTTGACTCTACTGAGAGTTGTAACCAAGTATCAGAGGTGCTCATTCACACG 280
DB 233 AGAAGCATCTCACCTTCCTTCGAAATTCGCTCCAAAGAAATTAATGATGTCACACG 292
QY 281 ATGCCAAGATCACCGTTTGGTGGTGCAGATGCCAAGA 320
DB 293 ATAAAGACTATTTCTGATTTGTGATTCAGATCCACCGA 332

RESULT 14
ABQ60352/c
ID ABQ60352 standard; cDNA; 650 BP.
XX
AC ABQ60352;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:4047.
XX
DE Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KM genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
XX Homo sapiens.
XX OS
XX WO200229086-A2.
XX

PD 11-APR-2002.
XX 02-OCT-2001; 2001WO-US30732.
PF 02-OCT-2000; 2000US-237271P.
PR 02-OCT-2000; 2000US-237271P.
XX
XX (FARB) BAYER CORP.
XX Burgess C, Ascle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA,
PI Thiglingam A, Lewis ME;
XX WPI; 2002-426115/45.
XX
XX New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy -
XX
XX Claim 1; Fig 1; 796pp; English.
XX
XX ABO6306 to ABO60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABO60776 to ABO60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridizes to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate.
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling; forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.
XX
XX Sequence 650 BP; 139 A; 129 C; 175 G; 191 T; 16 other;
SQ
Query Match 15.0%; Score 51.8; DB 24; Length 650;
Best Local Similarity 47.9%; Pred. No. 2.9e-07;
Matches 134; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 41 ACGTGAAGAAGTATTCGCAAGTCAGAGAAACCCGCGTGGAGACATATTGATCA 100
DB 638 AAGGTGAAGAAACANTGAAGNAGTGCAGGCCAGAGAGAGTGCAGGATTMAATGTCG 579
QY 101 TGAATCACTCGGGTGTCCGGTGAAGAACTCGATGATGCTCAGAGGGCTTGCACTAGC 160
DB 578 TGAACACAGAAAGGCAATTTCATCAGAGCAACATGNAACAACCCACACCCAGTATG 519
QY 161 CCTGTATATATGACATTTGGCGGAGAGATGCCAGCGCTTCTCCAAATATGAGCCAG 220
DB 518 CCAGCCTTCAGCACTTCATCTTAAGGACGAGAGCAACCGTGCCTGATGACCCCC 459
QY 221 CCCAAATTTGACTACTGAGAGTTGCTACCAAGTATCAGAGTGTCTATTACACAG 280
DB 458 AGAACATCTCACCTTCCTTGAATTCGCTCCAAAGAAAATGAATTAATGTTGCACAG 399
QY 281 ATGCCAAGATCAACGTTTGGTGTTCAGAAATGCCAAGA 320
DB 398 ATTAAGACTATTTCTGATTTGTGATTCAGATTCACCAAG 359

KW Human; km23; transforming growth factor-beta; TGF-beta;
KW signal transduction; ovary cancer; tumour suppressor; diagnosis;
KW gene therapy; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 119..409
FT /tag= a
FT /transl_except= "(pos:114..136,aa:Glu)"
FT /transl_except= "(pos:158..160,aa:Glu)"
FT /transl_except= "(pos:332..334,aa:Arg)"
XX
XX WO200162791-A2.
XX
XX 30-ANG-2001.
XX
XX 26-FEB-2001; 2001WO-US06176.
XX
XX 25-FEB-2000; 2000US-0184843.
XX 23-OCT-2000; 2000US-0242464.
XX (MULD/) MUDLER K M.
XX
XX Mulder KM;
XX
XX WPI: 2001-557699/62.
XX P-PSDB; AAB82823.
XX
XX New km23 polypeptide, a mediator of signal transduction activity of
PT transforming growth factor beta superfamily members, for diagnosing,
PT preventing and treating diseases related with km23 expression, e.g.
PT cancer -
XX
XX Disclosure; Fig 1A-B; 117pp; English.
XX
XX The present sequence is that of cDNA encoding human km23 (see
CC AAB82823), a dynein-associated protein and signalling intermediate
CC in the transforming growth factor-beta (TGF-beta) signal
CC transduction pathway. The cDNA was obtained from a placental
CC expression library. Human km23 protein interacts with TGF-beta
CC receptors, is phosphorylated by them, and synergizes with TGF-beta
CC to stimulate transcriptional activation of a cAMP-responsive
CC element reporter. Mutations in the oxido-reductase, protein kinase
CC C and casein kinase II phosphorylation sites, and also in the
CC Kringle domain of km23 are associated with ovarian tumours,
CC suggesting that the km23 gene acts as a tumour suppressor gene.
CC The invention provides pharmaceutical compositions including km23
CC polypeptides, fragments and mutants, polynucleotides encoding and
CC identifying km23, oligonucleotides and primers, expression vectors,
CC host cells, agonists, antibodies, antagonists and their use for the
CC diagnosis, prevention and treatment of diseases associated with
CC km23 expression or activity, or with defects in the signalling
CC pathway for TGF-beta superfamily members, especially for modulating
CC regulation of gene expression, cell growth, cell differentiation,
CC cell survival, apoptosis, senescence, cell migration, angiogenesis,
CC fibrosis, wound healing, extracellular matrix induction, adhesion,
CC autophagocytosis, and embryogenesis. Mutations in the km23 gene can
CC be used for diagnosis and prognosis of cancer, especially ovarian
CC cancer.
XX
XX Sequence 763 BP; 212 A; 219 C; 171 G; 161 T; 0 other;
SQ
Query Match 14.8%; Score 51.2; DB 22; Length 763;
Best Local Similarity 48.9%; Pred. No. 5e-07;
Matches 137; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 41 ACGTGAAGAAGTATTTGCGCAAGTCAGAGAAACCCGCGTGGAGACATATTGATCA 100
DB 126 AAGTGAAGACACACTGAAGCACTGCAGAGCCAGAGAGAGTGCAGAAATCATCTGTCG 185
QY 101 TGAATCACTCGGGTGTCCGGTGAAGAACTCGATGATGCTCAGAGGGCTTGCACTAGC 160

```

Db      186 TGAACACGAGAGGCAATCCCATCAAGAGCAACATGACAAACCCACCAACCAAGTATG 245
Qy      161 CTTGTCTATATGCAATTTGGGGAGAGTCCAGGGGTTCTCTCCAAATGAGGCGAG 220
Db      246 CCGACCTCATGACACAGCTTCATCTCGAAGGACGAGACCGTGCCTGACATGACCC 305
Qy      221 CCCAAATTTGACTCTACTGAGAGTTCGTACCAAGTATCAAGAGTGTCTCATTAACACG 280
Db      306 AGAAGATCTCACCCTTCCTCGAATTAAGTCCAAAGAAATGAATATATGTTGCACACG 365
Qy      281 ATGCCAAGATCACCGTTTGGTGTTCGAATGCCAAGA 320
Db      366 ATAAAGACTATTCTCGATTGTGATTCAGATCCACCGA 405

```

Search completed: February 7, 2004, 19:25:29
 Job time : 101.295 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:37:05 ; Search time 21.4835 Seconds
(without alignments)
7088.094 Million cell updates/sec

Title: US-09-614-150A-2

Perfect score: 345
Sequence: 1 atgcagagacagacagacaga.....tcttaacataaagggttag 345

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/pcodata/2/ina/5A COMB.seq:*
2: /cgn2_6/pcodata/2/ina/5B COMB.seq:*
3: /cgn2_6/pcodata/2/ina/6A COMB.seq:*
4: /cgn2_6/pcodata/2/ina/6B COMB.seq:*
5: /cgn2_6/pcodata/2/ina/PCETUS COMB.seq:*
6: /cgn2_6/pcodata/2/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	34.2	9.9	7218	1	US-08-232-463-14 Sequence 14, Appl
2	32.6	9.4	3027	2	US-08-680-326-23 Sequence 23, Appl
3	31.6	9.2	5056	2	US-08-793-126-2 Sequence 2, Appl
4	31.6	9.2	5056	3	US-09-132-271-2 Sequence 2, Appl
5	31.6	9.2	5057	3	US-09-142-334-23 Sequence 23, Appl
6	31.4	9.1	933	4	US-09-134-001C-2350 Sequence 2350, Ap
7	30.8	8.9	1158	4	US-09-252-991A-8509 Sequence 8509, Ap
C 8	30.8	8.9	2019	4	US-09-252-991A-8334 Sequence 8334, Ap
9	30.8	8.9	2703	4	US-09-252-991A-8440 Sequence 8440, Ap
10	30.8	8.9	4108	3	US-08-981-729-8 Sequence 8, Appl
11	30.8	8.9	4108	4	US-08-981-446B-1 Sequence 1, Appl
C 12	30.4	8.8	1590	4	US-09-252-991A-13186 Sequence 13186, A
13	30.4	8.8	1749	4	US-09-252-991A-12770 Sequence 12770, A
C 14	30	8.7	837	4	US-09-252-991A-8260 Sequence 8260, Ap
15	30	8.7	1434	4	US-09-252-991A-8003 Sequence 8003, Ap
C 16	30	8.7	2403	4	US-09-252-991A-8064 Sequence 8064, Ap
17	29.6	8.6	1709	2	US-09-010-398-2 Sequence 2, Appl
18	29.6	8.6	1709	3	US-09-366-260-2 Sequence 2, Appl
C 19	29.6	8.6	1709	3	US-09-153-804-11 Sequence 11, Appl
20	29.4	8.5	4935	2	US-08-631-097-3 Sequence 3, Appl
C 21	29.4	8.5	5886	3	US-08-810-712-9 Sequence 9, Appl
22	29.2	8.5	606	3	US-09-305-640-3 Sequence 3, Appl
23	29.2	8.5	606	4	US-09-280-116-121 Sequence 121, App
24	29.2	8.5	1956	4	US-08-559-896B-1 Sequence 1, Appl
25	29.2	8.5	2560	3	US-09-305-640-1 Sequence 463, App
26	29.2	8.5	2853	4	US-09-620-312D-463 Sequence 7, Appl
27	29.2	8.5	9381	4	US-09-453-702B-7

C 28	29.2	8.5	49377	1	US-08-764-233A-1 Sequence 1, Appl
29	28.8	8.3	1332	4	US-09-252-991A-2204 Sequence 2204, Ap
30	28.8	8.3	3335	4	US-09-252-991A-2282 Sequence 2282, Ap
31	28.6	8.3	289	3	US-09-007-005-17 Sequence 17, Appl
32	28.6	8.3	289	3	US-09-244-796-17 Sequence 17, Appl
33	28.4	8.2	474	6	5514566-5 Patent No. 5514566
34	28.4	8.2	507	5	PCT-US91-02186-5 Sequence 5, Appl
35	28.4	8.2	1002	4	US-09-170-496D-15 Sequence 15, Appl
36	28.4	8.2	1002	4	US-09-170-496D-171 Sequence 171, App
37	28.4	8.2	1518	1	US-08-148-215A-3 Sequence 3, Appl
38	28.4	8.2	1518	4	US-09-016-434-1460 Sequence 1460, Ap
39	28.4	8.2	1869	4	US-09-328-352-3844 Sequence 3844, App
40	28.4	8.2	3250	4	US-09-328-352-3844 Sequence 960, App
C 41	28.4	8.2	34094	4	US-09-292-034-1 Sequence 1, Appl
42	28.2	8.2	318	4	US-09-252-991A-15823 Sequence 15823, A
C 43	28.2	8.2	915	4	US-09-252-991A-15932 Sequence 15932, A
44	28.2	8.2	1068	4	US-09-298-886-9 Sequence 9, Appl
45	28.2	8.2	1149	4	US-09-298-886-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONING: pTZ9pt-F1s
; US-08-232-463-14
Query Match 9.9%; Score 34.2; DB 1; Length 7218;

[illegible]

```

1      RESULT 2
2      US-08-680-326-23
3      : Sequence 23, Application US/08680326
4      : Patent No. 5925733
5      : GENERAL INFORMATION:
6      : APPLICANT: ROSE, TIMOTHY M.
7      : APPLICANT: BOSCH, MARINIX
8      : APPLICANT: STRAND, KURT
9      : APPLICANT: TODARO, GEORGE J.
10     : TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES
11     : TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL
12     : TITLE OF INVENTION: FIBROMATOSIS
13     : NUMBER OF SEQUENCES: 152
14     : CORRESPONDENCE ADDRESS:
15     : ADDRESSEE: MORRISON & FOERSTER
16     : STREET: 755 Page Mill Road
17     : CITY: Palo Alto
18     : STATE: California
19     : COUNTRY: USA
20     : ZIP: 94304-1018
21     : COMPUTER READABLE FORM:
22     : MEDIUM TYPE: floppy disk
23     : COMPUTER: IBM PC compatible
24     : OPERATING SYSTEM: PC-DOS/MS-DOS
25     : SOFTWARE: PatentIn Release #1.0, Version #1.30
26     : CURRENT APPLICATION DATA:
27     : APPLICATION NUMBER: US/08/680,326
28     : FILING DATE:
29     : CLASSIFICATION: 514
30     : ATTORNEY/AGENT INFORMATION:
31     : NAME: Schiff, J. Michael
32     : REGISTRATION NUMBER: 40,253
33     : REFERENCE/DOCKET NUMBER: 29938-20001.00
34     : TELECOMMUNICATION INFORMATION:
35     : TELEPHONE: (415) 813-5600
36     : TELEFAX: (415) 494-0792
37     : TELEX: 706141
38     : INFORMATION FOR SEQ ID NO: 23:
39     : SEQUENCE CHARACTERISTICS:
40     : LENGTH: 3027 base pairs
41     : TYPE: nucleic acid
42     : STRANDEDNESS: double
43     : TOPOLOGY: linear
44     :
45     : US-08-680-326-23
46
47     Query Match 9.4%; Score 32.6; DB 2; Length 3027;
48     Best local Similarity 47.3%; Pred. No. 0.4; Mismatches 109; Indels 0; Gaps 0
49     Match 98; Conservative 0;

```

QY	65	TGCAAGAGAAACCCGCGGTGAGAGCAATTATGATCATGAATTCACCTCGGGTGTGCGGTGA	124
Db	1826	TGCACCCGCACCTCGGGCCGGGCGCATACGAGACCTTTGAGCTCCGAGGGGGCGGTGC	1885
QY	125	AAACCTCGATGATCGTCAGAGAGGCTTGCAGTACGCTGTCTATATGACAAATTGCGGG	184
Db	1886	ACTTTGTCAAGAAACACAAAGCGGTCTCGCTGCTGGCCACGCTGTGAACGTGTGGCTGG	1945
QY	185	AGAAGTGCCAGGCGTTCCTCTCCAAATATGAGCCAGCCCAAAATTGATCTCTACTGAGAG	244
Db	1946	CCAAAGAGAAAGCCCATAGGCGCGAAGCTGGCCACGCTTCGAGACGAGGCGGTACAGACCA	2005
QY	245	TTGCTACCAAGTATCAGAGGTGCTCA	271
Db	2006	TCCTGAGACAGACGACGTGGCCATCA	2032

```

RESULT 3
US-08-793-126-2
Sequence 2, Application US/08793126
Patent No. 5849297
GENERAL INFORMATION:
APPLICANT: Harrison, Richard Alexander
APPLICANT: Farris, Charles Timothy
TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States of America
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,126
FILING DATE: 07-FEB-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Baker, Hollie L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 102286,377
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5056 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-793-126-2

Query Match          9.2%; Score 31.6; DB 2; Length 5056;
Best Local Similarity 46.7%; Pred. No. 1.1;
Matches 100; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

      81  CGTGAAGACATATGATGATCACTCGGGGTGCGGTAAGAAACCTGATGATGCG 140
      910  CCTAAGCGCAATTCGATTGAGATGCGTGGGGGAAGTTGTCTGACCGGAAGTACT 969
      141  TCAGAGAGCGCTTCGAGTAGCGCTGTCTATATGACAAATTTGGCGGAGAAATGCCAGCGT 200
      970  GCTGAGCGGGGTGCAAGACCCCGACGAGAAAGCCTGCTGGGGAAGCTTTGTACGTGTC 1029
      201  CCTTCCAAATGAGACCGCAACAAATTTGACTCTACTAGAGTGTGTAACCAAGTATCA 260
      1030  TGCACCGTCATCTTGCACTTAGCGACGTGACATGTGCAAGCAGACGCGGGAATCC 1089

```

QY 261 CGAGTGCTCATTAACCAAGATCCAGATCACC 294
DB 1090 CATCGTACCTCTCCACCAAGATCCACTTACC 1123

RESULT 4

US-09-132-271-2

; Sequence 2, Application US/09132271

; Patent No. 6221657

; GENERAL INFORMATION:

; APPLICANT: Harrison, Richard Alexander

; APPLICANT: Faries, Charles Timothy

; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HALE AND DORR LLP

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: United States of America

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/132,271

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/793,126

; FILING DATE: 07-FEB-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Baker, Hollie L.

; REGISTRATION NUMBER: 31,321

; REFERENCE/DOCKET NUMBER: 102286.377

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 526-6000

; TELEFAX: (617) 526-5000

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5056 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-09-132-271-2

Query Match 9.2%; Score 31.6; DB 3; Length 5056;
Best Local Similarity 46.7%; Pred. No. 1.1;

Matches 100; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 81 CGTGAGACATATTTGATGATCACTCGGGTGGCGGTAAGAAACCTCGATGATCG 140
DB 910 CTTCAAGCCGATTCGATGATGAGATGCTCGGGAGGTTGTCTGAGCCGGAAGTACT 969
QY 141 TCAGAGGGGCTTGACGATCGCTGTCTATATGACAAATTTGCGGAGAAAGTCCAGGCGTT 200
DB 970 GCTGAGCGGGGTGACAGAACCCCGAGCAGAAAGCTGTGGGAAAGTCTTTGTAAGGTGTC 1029
QY 201 CCTCTCCAAATATGAGACCCCAAAATTTGACTCTACTGAGAGTTGCTACCAAGTATCA 260
DB 1030 TGGCACCCTGATCTTGTGACTCAGGCAAGTGAATGTGACAGGAGAGCGGAGATCCC 1089
QY 261 CGAGTGCTCATTAACCAAGATCCAGATCACC 294
DB 1090 CATCGTACCTCTCCACCAAGATCCACTTACC 1123

RESULT 5
US-09-142-334-23

; Sequence 23, Application US/09142334

; Patent No. 6268485

; GENERAL INFORMATION:

; APPLICANT: Harrison, Timothy C.

; APPLICANT: Faries, Richard A.

; TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase

; FILE REFERENCE: 4-30443/A/IMU/PCT

; CURRENT APPLICATION NUMBER: US/09/142,334

; EARLIER FILING DATE: 1999-04-15

; EARLIER APPLICATION NUMBER: PCT/GB97/00603

; EARLIER FILING DATE: 1997-03-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: Patent Ver. 2.0

; SEQ ID NO 23

; LENGTH: 5067

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-142-334-23

Query Match 9.2%; Score 31.6; DB 3; Length 5067;
Best Local Similarity 46.7%; Pred. No. 1.1;

Matches 100; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 81 CGTGAGACATATTTGATGATCACTCGGGTGGCGGTAAGAAACCTCGATGATCG 140
DB 921 CTTCAAGCCGATTCGATGATGAGATGCTCGGGAGGTTGTCTGAGCCGGAAGTACT 980
QY 141 TCAGAGGGGCTTGACGATCGCTGTCTATATGACAAATTTGCGGAGAAAGTCCAGGCGTT 200
DB 981 GCTGAGCGGGGTGACAGAACCCCGAGCAGAAAGCTGTGGGAAAGTCTTTGTAAGGTGTC 1040
QY 201 CCTCTCCAAATATGAGACCCCAAAATTTGACTCTACTGAGAGTTGCTACCAAGTATCA 260
DB 1041 TGGCACCCTGATCTTGTGACTCAGGCAAGTGAATGTGACAGGAGAGCGGAGATCCC 1100
QY 261 CGAGTGCTCATTAACCAAGATCCAGATCACC 294
DB 1101 CATCGTACCTCTCCCTTACCAAGATCCACTTACC 1134

RESULT 6

US-09-134-001C-2350

; Sequence 2350, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stream et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 2350

; LENGTH: 933

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-2350

Query Match 9.1%; Score 31.4; DB 4; Length 933;
Best Local Similarity 47.7%; Pred. No. 0.58;

Matches 92; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 47 AAGAGTATTTGGCAAGTGAGAGAAACCCGGCGTGAGAGACATATGATGATGATC 106
DB 528 AAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 587
QY 107 ACTCGGCTGCGCGTGAAGAACCTCGATGATGATGATGATGATGATGATGATGATGAT 166
DB 588 AAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 647

OY	244	GTTCTACCAAGTATACGAGGTGCTCATTTACACGAGTCCAAATATACCGTTTGGTG	303
Db	1468	TTGCGCCGNGTATATCCGGGTGCTGACTTCCGGCAGTGGCGTGTCTCACTCGGTGAAA	1527
OY	304	GTTCCAGATG	313
Db	1528	GCCAGAAAG	1537

RESULT 10

```

US-08-981-729-8
: Sequence 8, Application US/08981729
: Patent No. 6177261
:
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: A novel method to isolate mutants and to
: TITLE OF INVENTION: clone the complementing gene
: NUMBER OF SEQUENCES: 10
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/981,729
: FILING DATE:
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4108 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Aspergillus niger (CBS 120.49)
: STRAIN: NM147
: FEATURE:
: NAME/KEY: TATA_signal
: LOCATION: 787..794
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 855..3266
: OTHER INFORMATION: /EC number= 3.2.1.37
: OTHER INFORMATION: /product= "1,4-beta-D-xylan xylohydrolase"
: OTHER INFORMATION: /gene= "xlnD"
: OTHER INFORMATION: /standard_name= "beta-xyloisidase"
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: 855..932
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 933..3266
: FEATURE:
: NAME/KEY: polyA_site
: LOCATION: 3383
: FEATURE:
: NAME/KEY: polyA_site
: LOCATION: 3404
:
US-08-981-729-8

```

	Query Match	8.9%	Score 30.8	DB 3	Length 4108		
	Best Local Similarity	51.4%	Pred. No. 2				
	Matches	71	Conservative	0	Mismatches 67	Indels 0	Gaps 0
Qy	60	CAANATGAGAGAAACCGCGCTGCGAGACATATTGATTCATCTCGGGGTGCC	119				
Db	261	CAMGAGCTGATTAATACGTGGGCGCCAGAGGTCAAGATCACGTTTGAGATGGGAGCT	320				
Qy	120	GGTGAACCTTCATGATTCGTACGAGAGCGCTTGCAATCGCTGTCTATATGACATTT	179				
Db	321	GGTGAAGAGTTGTGGAAGGCTCCGACAGAGGATCATCTTGGGGTGAGAAATGAGAGGT	380				

QY	180	GCGGGAAGTCCAGGC	197
Db	381	GTTGCAAGATGCGGCGC	398

RESULT 11

```

US-08-981-446B-1
: Sequence 1, Application US/08981446B
: Patent No. 6300112
:
: GENERAL INFORMATION:
:
: APPLICANT:
:
: TITLE OF INVENTION: No. 6300112el beta-xylosidase, nucleotide sequence
: TITLE OF INVENTION: encoding it, and use thereof
: NUMBER OF SEQUENCES: 3
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentn Release #1.0, Version #1.25 (EPO)
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/981,446B
:
: FILING DATE:
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4108 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEITICAL: NO
: ANTI-SENSE: NO
:
: ORIGINAL SOURCE:
: ORGANISM: Aspergillus niger (CBS 120.49)
: STRAIN: NM147
:
: FEATURE:
: NAME/KEY: TATA_signal
: LOCATION: 787..794
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 855..3266
: OTHER INFORMATION: /EC number= 3.2.1.37
: OTHER INFORMATION: /product= "1,4-beta-D-xylan xylohydrolase"
: OTHER INFORMATION: /gene= "xlnD"
: OTHER INFORMATION: /standard_name= "beta-xylosidase"
:
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: 855..932
:
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 933..3266
:
: FEATURE:
: NAME/KEY: polyA_site
: LOCATION: 3383
:
: FEATURE:
: NAME/KEY: polyA_site
: LOCATION: 3404
:
US-08-981-446B-1

```

	Query Match	Similarity	8.9%	Score 30.8	DB 4	Length 4108
	Best Local	Similarity	51.4%	Pred. No. 2		
	Matches	71	Conservative	0	Mismatches	67
					Indels	0
					Gaps	0
Oy	60	CAAAAGTCAGAGAAACCCGCGCTGAGAGACATATTGATCATGAATCATCTCGGGTGTCC				119
Db	261	CAAGGACTCGATATATCTGGGGCGCAGAGGCTCAAGATCATCGTTTGAGATGGGGAGGT				320
Oy	120	GGTAAAAACCTCCATGATGTGTCTGAGAGGGCTTGACAGTACGCTGTCTATATGACAATTT				179
Db	321	GGTGAAGAAGTGTGTTGAAGGCTCCGAAGGAGTCCATCTGGGGTGACGAATGAGAGGT				380
Oy	180	GGGGGAGAAAGTCCAGGC				197

Db 381 GTTCAGAGAGTGGCGGC 398

RESULT 12

US-09-252-991A-13186/c
; Sequence 13186, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13186
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13186

Query Match 8.8%; Score 30.4; DB 4; Length 1590;

Best Local Similarity 61.2%; Pred. No. 1.7;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 61 AAAGTCAGAGAAACCGCGGTGAGAGACATATTGATCATGATCACTCGGGTGTCCG 120

Db 599 AACCGCCGAGAAACCGAGCGGTGTGGCGAGGCTGTGCATCACCCCTTGGGCGCCCG 540

QY 121 GTGAAAACCTCGATGATCG 140

Db 539 GTGGAACCGAGGTGTATTG 520

RESULT 13

US-09-252-991A-12770
; Sequence 12770, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12770
; LENGTH: 1749
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12770

Query Match 8.8%; Score 30.4; DB 4; Length 1749;

Best Local Similarity 61.2%; Pred. No. 1.7;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 61 AAAGTCAGAGAAACCGCGGTGAGAGACATATTGATCATGATCACTCGGGTGTCCG 120

Db 1136 AACCGCCGAGAAACCGAGCGGTGTGGCGAGGCTGTGCATCACCCCTTGGGCGCCCG 1195

QY 121 GTGAAAACCTCGATGATCG 140

Db 1196 GTGGAACCGAGGTGTATTG 1215

RESULT 14

US-09-252-991A-8260
; Sequence 8260, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8260
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8260

Query Match 8.7%; Score 30; DB 4; Length 837;

Best Local Similarity 49.4%; Pred. No. 1.7;
Matches 78; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 67 CAGAGAAACCGCGGTGAGAGACATATTGATCATGATCACTCGGGTGTCCGAGAA 126

Db 205 CGGCGGACAGCGGGGTAGTCCGCTTGGAGCAGGCCAGGCGCGAT 264

QY 127 ACCTGATGATGTCAGAGAGGCTTGACATGCGCTGTCTATATGCAATTTGCGGAG 186

Db 265 ATCGGACACACCGGCGCTTCAGACGCGACCGGCTGCTGTGGCCATGTTGCCGCG 324

QY 187 AAGTCCAGCGCTTCTCTTCCAAATGAGCCAGCCCA 224

Db 325 ACGTACCAAGTGTCTCTCGCGTATTGATGCGCGCA 362

RESULT 15

US-09-252-991A-8003/c
; Sequence 8003, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8003
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8003

Query Match 8.7%; Score 30; DB 4; Length 1434;

Best Local Similarity 49.4%; Pred. No. 2.2;
Matches 78; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 67 CAGAGAAACCGCGGTGAGAGACATATTGATCATGATCACTCGGGTGTCCGAGAA 126

Db 901 CGGCGGACAGCGGGGTAGTCCGCTTGGAGCAGGCCAGGCGCTGCGGAT 842

QY 127 ACCTGATGATGTCAGAGAGGCTTGACATGCGCTGTCTATATGCAATTTGCGGAG 186

Db 841 ATCGGACACACCGGCGCTTCAGACGCGACCGGCTGCTGTGGCCATGTTGCCGCG 782

Qy 187 AAGTCCAGGCGTTCTCTCCAAATGAGCCAGCCCA 224
| | | | | | | | | | | | | | | | | | | | | |
Db 781 ACGTACCAAGGTGGTCTCTGCGGTAAATTGATGCCGCCGA 744

Search completed: February 8, 2004, 10:31:05
Job time : 24.4835 secs

THIS PAGE BLANK (USPTO)

Db 128 CCAGCCTCATGACAGCTTCCTCTGAAGCAGGACCGCTGCTGACATCGACCC 187
Qy 221 CCCAAATTGACTCTCTAGAGTTGTAACAGATTCAGAGGTCTCATTAACACG 280
Db 188 AGAACGATCTCAGCTTCTCTGAAATTCGCTCAAGAAATTAATTAATGATGACACG 247
Qy 281 ATGCCAAGATCACCGTTTGGTGGTTCAGATGCCAAGA 320
Db 248 ATAAAGACTATTTCTGATTTGTGATTCAGATCCACGA 287

RESULT 2
US-09-941-831-10
; Sequence 10, Application US/09941831
; Patent No. US20020160493A1
; GENERAL INFORMATION:
; APPLICANT: Ebner et al.
; TITLE OF INVENTION: PT049P1
; FILE REFERENCE: Serine/Threonine Phosphatase Polynucleotides, Polypeptides, and Antih
; CURRENT APPLICATION NUMBER: US/09/941,831
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: PCT/US01/06256
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/186,350
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-941-831-10

Query Match 15.3%; Score 52.8; DB 10; Length 665;
Best Local Similarity 49.3%; Pred. No. 1.8e-07;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy 41 ACCTGGAAGAATGATTTTCGAAAGTGCAGAGAAACCCGGCTGAGACATATTGATCA 100
Db 31 AGGTGAGAGACACTGAGAGCTGCAGAGCCAGAGAGTGCAGAGATCTGCTCG 90
Qy 101 TGAATCACTCGGGTGTCCGGTGAAGAACTTCGATGATCTGTCAGAGAGGCTTGCATGAC 160
Db 91 TGAACACAGAAAGGATTCCTCATCAAGACACATGACCAACCCACCAACCCAGATATG 150
Qy 161 CCTGTCTATATGACAAATTTGGCGGAGAGTCCAGGCGTCTCTCCAAATGAGCCAG 220
Db 151 CCAGCCTCATGACAGCTTCATCTCTGAAGGACGAGACCGTGCATGACATGACCC 210
Qy 221 CCCAAATTGACTCTCTAGAGTTGTAACAGATTCAGAGGTCTCATTAACACG 280
Db 211 AGAACGATCTCAGCTTCTCTGAAATTCGCTCAAGAAATTAATTAATGATGACACG 270
Qy 281 ATGCCAAGATCACCGTTTGGTGGTTCAGATGCCAAGA 320
Db 271 ATAAAGACTATTTCTGATTTGTGATTCAGATCCACGA 310

RESULT 3
US-10-227-988A-3
; Sequence 3, Application US/10227988A
; Publication No. US20030148322A1
; GENERAL INFORMATION:
; APPLICANT: Mulder, Kathleen M.
; TITLE OF INVENTION: Control of TGF β Growth Factor Signaling by km23 and its Derivativ
; FILE REFERENCE: 1151-3 CIP
; CURRENT APPLICATION NUMBER: US/10/227,988A
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: U.S. Provisional Application Serial No. US20030148322A1 60/
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: U.S. Provisional Application Serial No. US20030148322A1 60/
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: International Application No. US20030148322A1 PCT/US01/0661

; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-227-988A-3

Query Match 15.3%; Score 52.8; DB 13; Length 721;
Best Local Similarity 49.3%; Pred. No. 1.9e-07;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy 41 ACCTGGAAGAATGATTTTCGAAAGTGCAGAGAAACCCGGCTGAGACATATTGATCA 100
Db 88 AGGTGAGAGACACTGAGAGCTGCAGAGCCAGAGAGTGCAGAGATCATGCTG 147
Qy 101 TGAATCACTCGGGTGTCCGGTGAAGAACTTCGATGATCTGTCAGAGGCTTGCATGAC 160
Db 148 TGAACACAGAAAGGATTCCTCATCAAGACCATGAGCAACCCACCAACCCAGATATG 207
Qy 161 CCTGTCTATATGACAAATTTGGCGGAGAGTCCAGGCGTCTCTCCAAATGAGCCAG 220
Db 208 CCAGCCTCATGACAGCTTCATCTCTGAAGGACGAGACCGTGCATGACATGACCC 267
Qy 221 CCCAAATTGACTCTCTAGAGTTGTAACAGATTCAGAGGTCTCATTAACACG 280
Db 268 AGAACGATCTCAGCTTCTCTGAAATTCGCTCAAGAAATTAATTAATGATGACACG 327
Qy 281 ATGCCAAGATCACCGTTTGGTGGTTCAGATGCCAAGA 320
Db 328 ATAAAGACTATTTCTGATTTGTGATTCAGATCCACGA 367

RESULT 4
US-09-925-302-183
; Sequence 183, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 183
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (840)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (841)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (850)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-183

Query Match 15.3%; Score 52.8; DB 9; Length 858;
Best Local Similarity 49.3%; Pred. No. 2e-07;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy 41 ACCTGGAAGAATGATTTTCGAAAGTGCAGAGAAACCCGGCTGAGACATATTGATCA 100
Db 197 AGGTGAGAGACACTGAGAGCTGCAGAGCCAGAGAGGAGTGCAGAGATCATGCTG 256

```
QY 101 TGAATCACTCGGATGTCGCCGATGAAACCTGATGATCTGACGAGGCGCTTGACATACG 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 TGAACACAGAAAGGATTCATCAAGACGACATGACAAACCCACACACCAAGATG 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 161 CCTGTCTATATGACATTTGCGGAGAGATGCCAGCGCTTCTCTCCAAATGAGCCAG 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 CCAGCCTCATGACACAGCTTCATCTGTAAGGACGAGACACCGTGCATGACACCCCC 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 CCCAAATTGACTCTACTGAGAGTTGTCACAAAGTATCGAGAGTCTCATTAACCCAG 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 AGAACGATCTCACCTTCTTGAAATCGCTCCAGAAATGAAATTAATGTTGACACG 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 ATGCCAAGATCACCGTTTGTGCTTGCAGATGCGCAAGA 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 ATAAAGACTATTTCTGATTTGATTCAGATCCAAACCGA 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 5

```
US-09-825-294-52
; Sequence 52, Application US/09825294
; Patent No. US20020004491A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolck, John A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Piling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; CURRENT APPLICATION NUMBER: US/09/825,294
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-825-294-52
```

```
Query Match 15.2%; Score 52.4; DB 9; Length 396;
Best Local Similarity 48.9%; Pred. No. 1.9e-07;
Matches 137; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
```

```
QY 41 ACGTGAAGAAGTATTCGCAAGTGCAGAGAAACCCGCGTGCAGACATATTGATCA 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 AGGTGAGAGAGACATGACGACTGCANAGCCGAAAGGATGCGAGATCATCTGCG 112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 101 TGAATCACTCGGATGTCGCCGATGAAACCTGATGATCTGACGAGGCGCTTGACATACG 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 TGAACACAGAAAGGATTCATCAAGACGACATGACAAACCCACACCAAGATG 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 161 CCTGTCTATATGACATTTGCGGAGAGATGCCAGCGCTTCTCTCCAAATGAGCCAG 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 CCAGCCTCATGACACAGCTTCATCTGTAAGGACGAGACACCGTGCATGACACCCCC 232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 CCCAAATTGACTCTACTGAGAGTTGTCACAAAGTATCGAGAGTCTCATTAACCCAG 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 AGAACGATCTCACCTTCTTGAAATCGCTCCAGAAATGAAATTAATGTTGACACG 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 ATGCCAAGATCACCGTTTGTGCTTGCAGATGCGCAAGA 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 ATAAAGACTATTTCTGATTTGATTCAGATCCAAACCGA 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 6
US-09-970-966-52
; Sequence 52, Application US/09970966
; Patent No. US20020173638A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Stolck, John A.
; APPLICANT: Mollesh, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 81, 189
; OTHER INFORMATION: n = A,T,C or G
US-09-970-966-52
```

```
Query Match 15.2%; Score 52.4; DB 10; Length 396;
Best Local Similarity 48.9%; Pred. No. 1.9e-07;
Matches 137; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
```

```
QY 41 ACGTGAAGAAGTATTCGCAAGTGCAGAGAAACCCGCGTGCAGACATATTGATCA 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 AGGTGAGAGAGACATGACGACTGCANAGCCGAAAGGATGCGAGATCATCTGCG 112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 101 TGAATCACTCGGATGTCGCCGATGAAACCTGATGATCTGACGAGGCGCTTGACATACG 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 TGAACACAGAAAGGATTCATCAAGACGACATGACAAACCCACACCAAGATG 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 161 CCTGTCTATATGACATTTGCGGAGAGATGCCAGCGCTTCTCTCCAAATGAGCCAG 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 CCAGCCTCATGACACAGCTTCATCTGTAAGGACGAGACACCGTGCATGACACCC 232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 CCCAAATTGACTCTACTGAGAGTTGTCACAAAGTATCGAGAGTCTCATTAACCCAG 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 AGAACGATCTCACCTTCTTGAAATCGCTCCAGAAATGAAATTAATGTTGACACG 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 ATGCCAAGATCACCGTTTGTGCTTGCAGATGCGCAAGA 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 ATAAAGACTATTTCTGATTTGATTCAGATCCAAACCGA 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 7

```
US-10-369-186-52
; Sequence 52, Application US/10369186
; Publication No. US20030232056A1
; GENERAL INFORMATION:
; APPLICANT: Fanger, Gary R.
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; CURRENT APPLICATION NUMBER: US/10/369,186
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 293
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 81, 189
; OTHER INFORMATION: n = A,T,C or G
US-10-369-186-52
```

```
Query Match 15.2%; Score 52.4; DB 12; Length 396;
Best Local Similarity 48.9%; Pred. No. 1.9e-07;
```


; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 382
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-472-382

Query Match 9.6%; Score 33.2; DB 13; Length 473;
Best Local Similarity 12.1%; Pred. No. 0.88;
Matches 36; Conservative 79; Mismatches 183; Indels 0; Gaps 0;

QY 26 AGGGAAGTAAAGTTAGTGAAGAGATTTTGCAGAAAGTGCAGAGAGAAACCGCGCTGG 85
DB 57 AASORIFLHGNRISHVPAASFRACRNLTIIMLSNVLAIDMAAFTGLALLEQDLISDNA 116
QY 86 AGGACATATTGATCATGATCACTCGGATGCGCGTGAACCTCGATGATGTCAGG 145
DB 117 QLSVDPATFHGRLHTLHLDRCGLDGLFRGLALQYLYQDNALQALPDTFRD 176
QY 146 AGGAGCTTGACATAGCGCTGTCTATATGACAATTTGCGGAGAAAGTGCAGGCGTTCCT 205
DB 177 LGNLTFLHGNRISVPERAFRGLHSLDRLHQNRAVHVHPIAFRDLGRLMTLYLFAN 236
QY 206 CCAAAATGAGCCAGCCCAAAATTTGACTGAGAGTTGCTACCAAGTATCAGAGG 265
DB 237 NLSALPTEALAPLALQYLRINDPWCDCRARPMLWMLQKFRSSSEVPCLPQRLAGR 296
QY 266 TGCTCATTCACCAAGATCAAGATCACCGTTTGTGTTGTTGAGATGCCAAGATAC 323
DB 297 DLKRLANDLQGCATATGYPHPIWTGRATDEBPGLPKCCQPDAAADKASVLEGRPAS 354

Search completed: February 8, 2004, 20:44:10
Job time : 135.967 secs

THIS PAGE BLANK (USPTO)

/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="res0744"
/tissue_type="testis"
/dev_stage="spinning stage"
/clone_lib="less"

BASE COUNT 201 a 92 c 106 g 199 t 2 others

Query Match 16.8%; Score 58; DB 12; Length 600;
Best Local Similarity 49.8%; Pred. No. 1,1e-06;
Matches 145; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

ORIGIN

37 AGTTACGTGGAAGATATTTCGCAAGTGAGAGAAACCCGCGTGAGAGACATATG 96
225 AATGAACTGAGACACAAATTCGAAACTGTCAGTCACAAAGCGTGTGGCAGCTC 284
97 ATCATGATCATCTCGGTGTGCGCGTGAACCTCGATGATCTGACAGAGGCTTGCAG 156
285 GTTGTAAACGAGAAAGAAATACCGATAAACCACTTAAATATCATCATTTCCGTACA 344
157 TAGCGCTGTATATGCAATTTCCGAGAGAGTCCAGGCGTTCCTCCAAATGAG 216
345 TAGCGCGGCTATTTGACCGCGTGTGACAAAGCAAACTGTGCGAGACCTGAT 404
217 CCAAGCCCAAAATTGACTCTACTAGAGATTGTCACAAATGATCAGAGTCTCATTA 276
405 CCGCTTAATGACTTGAACATTTTGAAGTAAAGCAAAAGCAAGTATATGTCGCG 464
277 CCAAGATCCCAAGATCACCGTTTGTGTTGTCAGATGCCAAAGTACTTTT 327
465 CCGATTAAGAAATTTATATTAATAGTAGTTCAAAAATCTGTGATGATTTT 515

Db

Qy

RESULT 6
BI481570 539 bp mRNA linear EST 27-AUG-2001
LOCUS R664145.5prtime RE Drosophila melanogaster normalized Embryo pfic-1
DEFINITION Drosophila melanogaster cDNA clone R664145.5 similar to robl:
F8an0010751 GO: (axonemal dynein (GO:0005858); motor protein
(GO:0003774)) located on: 2R 54C3-54C3.; 05/16/2001, mRNA sequence.
BI481570
VERSION BI481570.1 GI:15316586
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 539)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson,
J., Champagne, M., Chavez, C., Dorsett, V., Fattah, D., Fritse, E., George,
R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mitera, S.,
Mungall, C. J., Nuno, J., Paclob, J., Paragas, V., Park, G.,
Poulsen, Vong, S., Wan, K., Yu, C., Lewis, S. E., Celisner, S. and Rubin,
G. M.
BDGP/HMT RE Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Based upon one or more reads of this clone where vector sequence
was present at both ends, this clone has been determined to contain
a cDNA insert on the order of 600-1000 bases. hit genomic
AB003803: arm:2R [12218360,12500926] estimated-cyto:54B8-54C11:
05/16/2001
Plate: RE:641 row: D column: 9
High quality sequence stop: 460.
Location/Qualifiers
1. .539

FEATURES
Source

/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE64145"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pfic-1"

BASE COUNT 181 a 108 c 115 g 134 t 1 others

Query Match 16.7%; Score 57.6; DB 12; Length 539;
Best Local Similarity 50.4%; Pred. No. 1,4e-06;
Matches 141; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

ORIGIN

41 AGGTGGAAGATATTTCGCAAGTGAGAGAAACCCGCGTGAGAGACATATGATCA 100
108 AGGTGGAAGATATTTCGCAAGTGAGAGAAACCCGCGTGAGAGACATATGATG 167
101 TGAATCATCTCGGTGTGCGCGTGAACCTCGATGATCTGACAGAGGCTTGCAGTAC 160
168 TCAACATGAAAGTATTCGGGTCAAAATCCAGCTGAGCAACATCACCGTTCAGTAC 227
161 CTTGTCTATATGCAATTTCCGAGAGAGTCCAGGCGTTCCTCCAAATGAGACG 220
228 CTGCGCTTAAGTATGATGCTGCGCGAGCAAGGCAAGTGTGTGAGGACTTGATCTT 287
221 CCAAAATTTGACTCTACTAGAGATTGTCACAAATGATCAGAGTCTCATTAACAG 280
288 CAAACGATGATCTTCTCGCGTGTGATGATGATGATGATGATGATGATGATGATG 347
281 ATGCCAAGATCACCGTTTGTGTTGTCAGATGCCAAGA 320
348 ACAAGACTTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 387

Db

Qy

RESULT 7
A1061910 557 bp mRNA linear EST 23-APR-2001
LOCUS LD34974.5prtime LD Drosophila melanogaster embryo p072 Drosophila
DEFINITION melanogaster cDNA clone LD34974.5 similar to robl: F8an0010751
"motor protein" located on: 2R 54C3-54C3.; 04/10/2001, mRNA
sequence.
A1061910
VERSION A1061910.2 GI:13769096
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 557)
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G. M.
BDGP/HMT Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AB003803: arm:2R [12218360,12500926]
estimated-cyto:54B8-54C11: 04/10/2001
Plate: LD:349 row: G column: 2
High quality sequence stop: 528

FEATURES
Source

FEATURES
source

POLYA=No.

Location/Qualifiers

1.557

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="LD34974"

/sex="male and female"

/dev_stage="0 to 24 hours mixed stage embryonic"

/lab_host="XJ1 Blue"

/clone_lib="LD Drosophila melanogaster embryo pot2"

/note="Organ: embryo; Vector: pot2; Site:1; EcorI; Site:2;

XhoI; Sized fractionated cDNAs were directly ligated into

pot2."

BASE COUNT

203 a 103 c 119 g 132 t

ORIGIN

Query Match

Best Local Similarity 50.4%; Pred. No. 1.4e-06;

Matches 141; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY

41 ACCTGGAAGATGTTTCGCAAGTGCAGAGAAACCCGCGTGCAGACATATTGATCA 100

DB

80 AGGTGAGAGAACCTCAAGAGATCCAGACCACAAAGGTGTGTGGTACATTTGTGG 139

QY

101 TGAATCACTCGGCTGTCGCGTGAACCTCGATGATGTCAGAGGCGTTGCAGTACG 160

DB

140 TCACACATGAGAGTATTCGCTCAATCCAGCTGCAGACATACCAACCGTTCACTAG 199

QY

161 CCTGTATATGACAAATTTGCGGAGAGAGTCCAGGCGTTCTCTCCAAATGAGACCG 220

DB

200 CTGGCTTAATGAGTCACTGCGGACGACGAGGACGAAATGTTGAGGAGCTTGATCTT 259

QY

221 CCCAAATTTGACTCTAGAGAGTTCGACCAAGTATCCAGAGTCTTATTAACACCG 280

DB

260 CCAACGACATGATATTCGCGGAGTGCATCAAGAGCAGATCATGTGTGACACCG 319

QY

281 ATGCCAAGATCACCGTTTGTGTTCCAGATGCAAGCA 320

DB

320 ACAGAGCTTATCTGATTTGATTCATCCAAACCCACCA 359

RESULT 8

BM364049

LOCUS

BS320036000G9 Subtracted Lewin Cattle Spleen Bos taurus cDNA clone

DEFINITION

BS320036000G9 5', mRNA sequence.

ACCESSION

BM364049

VERSION

BM364049.1 GI:18107418

KEYWORDS

EST.

SOURCE

Bos taurus (cow)

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 455)

Subtracted Lewin Cattle Spleen ESTs

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

W. M. Keck Center for Comparative and Functional Genomics

University of Illinois at Urbana-Champaign

340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL

61801, USA

Tel: 217 333 5998

Fax: 217 244 5617

Email: h-lewin@uiuc.edu

Funding for Cattle EST sequencing was provided by the USDA National

Research Initiative, Project No. 98-35205-6644, and a grant from

the Japanese Ministry of Agriculture Fisheries and Forestry to

H.A. Lewin and J.E. Womack. Base-calling/Quality scores: PHRED form

Washington University Genome Center. Vector-Trimming: Cross Match

from Washington University Genome Center PHRAP suite. This sequence

FEATURES
source

is vector free and at least 200bp in length.
Insert Length: 455 Std Error: 0.00
Plate: BS320036000 row: G column: 9
Seq primer: CCGCAGCTCGAAATTAACC
High quality sequence stop: 455.
Location/Qualifiers

1.455

/organism="Bos taurus"

/mol_type="mRNA"

/strain="Rangus"

/db_xref="taxon:9913"

/clone="BS320036000G9"

/sex="female"

/dev_stage="Adult"

/clone_lib="Subtracted Lewin Cattle Spleen"

/note="Organ: Spleen; Vector: pBlueScript SK(+/-); Site:1;

EcorI; Site:2; XhoI; This BS cDNA library was generated by

subtraction of the original non-normalized bovine spleen

library with 16,800 previously sequenced clones from a

bovine placenta cDNA library. The original library was

constructed as described by Band et al (2000), Genome

Research 10(9): 1359-1368."

BASE COUNT

123 a 126 c 111 g 93 t 2 others

ORIGIN

Query Match

Best Local Similarity 50.0%; Pred. No. 1.7e-06;

Matches 140; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY

41 ACCTGGAAGATGTTTCGCAAGTGCAGAGAAACCCGCGTGCAGACATATTGATCA 100

DB

46 AGGTGAGAGAACCTCAAGAGATCCAGACCACAAAGGTGTGTGGTACATTTGTGG 105

QY

101 TGAATCACTCGGCTGTCGCGTGAACCTCGATGATGTCAGAGGCGTTGCAGTACG 160

DB

106 TAAACACAGAGGNTTCCATCAACAGACCAATGACATCCACACACACATAG 165

QY

161 CCTGTATATGACAAATTTGCGGAGAGAGTCCAGGCGTTCTCTCCAAATGAGACCG 220

DB

166 CCACTTCACTGACCAATCTTATTTGAGGCGCGGACACCGTCCGCAATGACCCCC 225

QY

221 CCCAAATTTGACTCTAGAGAGTTCGACCAAGTATCCAGAGTCTTATTAACACCG 280

DB

226 AGAATGACCTCACTTCTTCTGATTCGCTCCAGAAATTAATGATGTTGCACAG 285

QY

281 ATGCCAAGATCACCGTTTGTGTTCCAGATGCAAGCA 320

DB

286 ATAAAGACTATTTCTGATTTGATTCAGATCACTGA 325

RESULT 9

BM362999

LOCUS

BS320051000H5 Subtracted Lewin Cattle Spleen Bos taurus cDNA clone

DEFINITION

BS320051000H5 5', mRNA sequence.

ACCESSION

BM362999

VERSION

BM362999.1 GI:18106368

KEYWORDS

EST.

SOURCE

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

W. M. Keck Center for Comparative and Functional Genomics

University of Illinois at Urbana-Champaign

340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL

61801, USA

Tel: 217 333 5998

Fax: 217 244 5617
Email: h-lewin@unc.edu
Funding for Cattle EST sequencing was provided by the USDA National Research Initiative, Project No. 98-35205-6644, and a grant from the Japanese Ministry of Agriculture Fisheries and Forestry to H.A. Lewin and J.E. Womack. Base-calling/Quality scores: PHRED form Washington University Genome Center. Vector-priming: Cross Match from Washington University Genome Center PHRAP suite. This sequence is vector free and at least 200bp in length.
Insert Length: 454 Std Error: 0.00
Plate: BS320051000 row: H column: 5
Seq primer: CGCCAGCTCGAATTAACCC
High quality sequence stop: 454.
Location/Qualifiers

FEATURES

source

1. 465
/organism="Bos taurus"
/mol_type="mRNA"
/strain="Angus"
/db_xref="taxon:9913"
/clone="BS32005100H5"
/sex="female"
/dev_stage="adult"
/clone_lib="Subtracted Lewin Cattle Spleen"
/note="Organ: Spleen; Vector: pBluescript SK(+/-); Site: 1: EcoRI; Site 2: XhoI; This BS cDNA library was generated by subtraction of the original non-normalized bovine spleen library with 16,800 previously sequenced clones from a bovine placenta cDNA library. The original library was constructed as described by Band et al (2000), Genome Research 10(9): 1359-1368."
BASE COUNT 124 a 127 c 110 g 93 t

ORIGIN

Query Match 16.2%; Score 56; DB 12; Length 454;
Best Local Similarity 50.0%; Pred. No. 3.9e-06;
Matches 140; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

41 ACCTGGAAGAAGTATTCGCAAGTCAGAGAAACCCGCGTGGAGACATATTGATCA 100
46 AGGTGAGAGAACCTTGAAGGACTCCAGACCCGAGAGGCTGACAGGATCATCTGG 105
101 TGAATCACTCGGGTGTGCGGTGAAAACCTCGATGATCGTCAGAGGGCTTGACATG 160
106 TAAACACAGAAAGGATTCCTCATCAAGACCATGACATCCACACACACAGATAG 165
161 CCGTCTATATGACAAATTTGGGAGAAAGTCCAGGGCTTCTCCAAATGAGCCAG 220
166 CCAACCTCATGACAACTTCATCTTGAAGGCCGAGAGCACCGTCCGGAATGACCCC 225
221 CCCAAATTTGACTCTACTGAGAGTGTGACCAAGTATCAGAGGTGCTCATTAACAC 280
226 AGAATGACCTCATTCTTGAATTTGCTCCAGAAAATGAATTAATTTGTTGACACG 285
281 ATGCCAAGATCACCGTTTGGTGTTCAGATGCCAAGA 320
286 ATAAAGCTATTCTGATTGTGATTCAGATCCACCTGA 325

RESULT 10

BI484967 465 bp mRNA linear EST 28-AUG-2001
LOCUS RE88166, sprime RE Drosophila melanogaster normalized Embryo pfic-1
DEFINITION Drosophila melanogaster cDNA clone RE88166 5 similar to robl:
F88010751 GO: (taxon:mal dynein (GO:0005858); motor protein
(GO:0003774)) located on: 2R 54C3-54C3:: 05/16/2001, mRNA sequence.

ACCESSION BI484967
VERSION BI484967.1 GI:15325159
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 465)
AUTHORS Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C.J., Nuno, J., Pacleb, J., Paragas, V., Park, S., Pounenavong, S., Wan, K., Yu, C., Lewis, S.E., Celnikier, S. and Rubin, G.M.
TITLE BDGP/HMI RE Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Based upon one or more reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases. hit genomic AE003803; arm:2R (12218360,12500926) estimated-cyto:54B8-54C11: 05/16/2001
Plate: RE.681 row: F column: 6
High quality sequence stop: 378.
Location/Qualifiers

FEATURES

source

1. 465
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE68166"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Iona"
/clone_lib="RE Drosophila melanogaster normalized Embryo pfic-1"
/note="Organ: embryo; Vector: pFic1; Site 1: XhoI; Site 2: BamHI. Library was kindly generated by Pietro Carninci at the RIKEN. The library was normalized and excised using Cte recombinase. Plasmid cDNA library."
BASE COUNT 142 a 104 c 106 g 113 t

ORIGIN

Query Match 16.2%; Score 56; DB 12; Length 465;
Best Local Similarity 50.0%; Pred. No. 3.9e-06;
Matches 140; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

41 ACCTGGAAGAAGTATTCGCAAGTCAGAGAAACCCGCGTGGAGACATATTGATCA 100
108 AGGTGAGAGAACCTTGAAGGACTCCAGACCCGAGAGGCTGACAGGATCATCTGG 167
101 TGAATCACTCGGGTGTGCGGTGAAAACCTCGATGATCGTCAGAGGGCTTGACATG 160
168 TCAACATGAAGGATTCCTGTAATCCACGCTGACCAACTACCAACCGTTCAATAG 227
161 CCGTCTATATGACAAATTTGGGAGAAAGTCCAGGGCTTCTCCAAATGAGGCCAG 220
228 CTGGCCTTAATGAGTCACTGCGGACAAAGGACAGAAAGTGTGAGGACTTGGATCTT 287
221 CCCAAATTTGACTCTACTGAGAGTGTGACCAAGTATCAGAGGTGCTCATTAACAC 280
288 CCAACGACATGACATTTCTCGGGTGGATTCAGAAAGCAGAGATCATGTGGCACCG 347
281 ATGCCAAGATCACCGTTTGGTGTTCAGATGCCAAGA 320
348 ACAAGACTTCATCTGATTGTATCCAAACCAACCGA 387

RESULT 11

BM107287 547 bp mRNA linear EST 21-NOV-2001
LOCUS 511471 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BM107287
ACCESSION BM107287.1 GI:17038357
VERSION BM107287.1 GI:17038357
KEYWORDS EST.
SOURCE Bos taurus (cow)

LOCUS	AVS89499	570 bp	mRNA	linear	EST 27-NOV-2001
DEFINITION	AVS89499 Bos taurus brain fetus Bos taurus cDNA clone E1BR006G10				
ACCESSION	AVS89499				
VERSION	AVS89499.1	GI:9700492			
KEYWORDS	EST.				
SOURCE	Bos taurus (cow)				
ORGANISM	Bos taurus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos. 1 (bases 1 to 570)				
AUTHORS	Takasuga,A., Hirotsune,S., Itoh,R., Itohzo,A., Suzuki,H., Aso,H. and Sugimoto,Y.				
TITLE	Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs				
JOURNAL	Nucleic Acids Res. 29 (122), E108 (2001)				
MEDLINE	21570554				
PMID	11713328				
COMMENT	Contact: Yoshikazu Sugimoto Animal Genetics Division Shirakawa Institute of Animal Genetics Oodaka, Niisigo, Nishi-shirakawa, Fukushima 961-8061, Japan Tel: 81-248-25-5641 Fax: 81-248-25-5725 Email: kazusugi@cocoa.ocn.ne.jp Single pass sequencing. This clone was obtained from a polyA-deleted cDNA library.				
FEATURES	Location/Qualifiers				
SOURCE	1..570				
	/organism="Bos taurus"				
	/mol_type="mRNA"				
	/db_xref="taxon:9913"				
	/clone="E1BR006G10"				
	/tissue_type="brain"				
	/dev_stage="fetus"				
	/lab_host="DH10B"				
	/clone_id="Bos taurus brain fetus"				
	/note="Vector: pZ11; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"				
BASE COUNT	141 a 167 c 137 g 125 t				
ORIGIN					
Query Match	16.2%; Score 56; DB 9; Length 570;				
Best Local Similarity	50.0%; Pred. No. 4.4e-06;				
Matches 140; Conservative 0; Mismatches 140; Indels 0; Gaps 0;					
41	ACGTGGAAGAGTATTTCGCAAGTCAGAGAAACCCGCGTGGAGCAATATTGATCA 100				
34	AGGTGGAGGAGACCCGTGAAGGACATCCAGACCAGAAAGGGCGTGCAGGAAATCATCTGG 93				
101	TGATTCACCTGGGGTGTGCGCGGGAACCTGATGATGATGTCAGAGAGGCTTTCAGATCG 160				
94	TAAACACGAAGGAGATTCCTCATCAAGACACACATGGAATATCCACCAACACAGATACG 153				
161	CGTGTCTATGACAAATTTGGCGGAGAGTGCAGGCGCTTCCTCCAAAATGAGAGCCAG 220				
154	CCAACCTATGACAACTTCATCTTGAAGGCCCGAGAGACCGTGGCGCAATATGACCCCC 213				
221	CCCAAAATTTGACTCTACTGAGAGTTGTCACCAAGTATCAGAGGTGCTCATTCACAGC 280				
214	AGAAATGACCTCACTTCCTCGAATTCGCTCCAGAAAAATGAATATATGTTGACAG 273				
281	ATGCCAAGATCACGTTTGGTGTGTCAGATGCCAAAG 320				
274	ATAAAGACTATTTCTGATGTGATTCAGAAATCAACTGA 313				

DEFINITION	AV589119 Bos taurus brain fetus Bos taurus cDNA clone E1BR004E07
ACCESSION	AV589119
VERSION	AV589119.1
KEYWORDS	EST.
SOURCE	Bos taurus (cow)
ORGANISM	Bos taurus
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovidae; Bos. 1 (bases 1 to 571)
AUTHORS	Takasuga, A., Hirota, S., Itoh, R., Ito, H., Suzuki, H., Abo, H., and Sugimoto, Y.
TITLE	Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
JOURNAL	Nucleic Acids Res. 29 (22), E108 (2001)
MEDLINE	21570554
PubMed	11713328
COMMENT	Contact: Yoshihiko Sugimoto Animal Genetics Division Shirakawa Institute of Animal Genetics Oadkura, Nishio, Aichi-shi, Japan Tel: 81-248-25-5641 Fax: 81-248-25-5725 Email: kazue@oac.oac.nie.jp Single pass sequencing. This clone was obtained from a polyA-deleted cDNA library. Location/Qualifiers 1. 571 /organism="Bos taurus" /mol_type="mRNA" /db_xref="taxon:9913" /clone="E1BR004E07" /issue_type="brain" /dev stage="fetus" /lab_host="DH10B" /clone_1ib="Bos taurus brain fetus" /note="Vector: pGL1; Site_1: SalI; Site_2: NotI; Poly A was deleted from a NotI site"
BASE COUNT	143 a 166 c 142 g 119 t 1 others
ORIGIN	
Query Match	16.2%; Score 56; DB 9; Length 571;
Best Local Similarity	50.0%; Pred. No. 4; de-06;
Matches 140; Conservative	0; Mismatches 140; Indels 0; Gaps 0;
QY	41 ACGGGAAGAAGTATTCGCAAGTGAAGGAAACCCGGGTGAGACATATTGATCA 100
DB	51 AGGGGAGGAGACCCCTTAAGCGATCCAGAGCCGAGAGGGGGTGCAGGATCTCGTG 110
QY	101 TGAATCACTGGGTGTCCGGTGAACCTCGATGATCGTCAGAGAGGGCTTGCAGTAGC 160
DB	111 TAAACAGAAAGGATTCCTCATCAAGAGCAACATGACCAATCCACACACACAGTAGC 170
QY	161 CCGTCTATATGACAAATTTGGGGGAGAGTCCAGGGCTTCTTCCAAATGAGCGAC 220
DB	171 CCAACCTCAGACACATTCATCTTGAAGGCCCGGACACCGTGGGAAATTGACCCC 230
QY	221 CCCAAATTTGACCTACTGAGAGTTGGTGAACAGATGACGAGGTGCTCATTCACACAG 280
DB	231 AGAATGACCTACTTCTTCGATTTGCTTCACAGAAATGAAATTAATGTTGTCACACAG 290
QY	281 ATGCCAAGATCACCGTTTGGTGGTTCAGATGACCAAGA 320
DB	291 ATAAAGACTATTTCTGATTTGTATTCAGATCACTGA 330
RESULT 15	
LOCUS	B0173798 504 bp mRNA linear EST 18-SEP-2002
DEFINITION	B0173798 B0173798.1 T7 Echinosuccus granulosis pepsin-treated proteoC1ex (trans-spliced) Echinosuccus granulosis cDNA clone

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Bg_PSPSL_02H12.5' similar to AAL75951 BITHORAXOID-LIKE PROTEIN.
Homo sapiens (Human), mRNA sequence.
B0173798
B0173798.2 GI:23187927
EST.

Echinococcus granulosus
Echinococcus granulosus
Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
Cyclophyllidae; Taeniidae; Echinococcus.
1 (bases 1 to 504)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Fernandez, C. and Maizels, R.M.
A survey of genes expressed in Echinococcus granulosus metacestode
Unpublished
On Apr 29, 2002 this sequence version replaced gi:20338200.
Contact: Maizels RM

Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, West Mains Road, Edinburgh EH93JT, Scotland
Tel: +44 131 650 5511
Fax: +44 131 650 5450

Email: r.maizels@ed.ac.uk

The library was created by Dr Cecilia Fernandez in conjunction with
Prof. R. Maizels. Sequencing was also performed by Cecilia.

PCR Primers

FORWARD: M13P

BACKWARD: M13R

Plate: 02 row: H column: 12

Seq primer: T7

High quality sequence stop: 468.

FEATURES
source

Location/Qualifiers
1..504

/organism="Echinococcus granulosus"

/mol_type="mRNA"

/db_xref="taxon:6210"

/clone="Bg_PSPSL_02H12"

/dev_stage="larva (protoscolex)"

/clone_1lb="Echinococcus granulosus pepsin-treated
protoscolex (trans-spliced)"

/note="Vector: pSPORT1; Site_1: SalI (5'end); Site_2: NotI
(3'end); Echinococcus granulosus is a cestode parasite of
dogs (definitive host) and various domestic and wild
animals as well as humans (intermediate hosts). The
library was constructed from protoscoleces obtained from
hydatid cysts present in the lungs of naturally infected
Uruguayan bovines. Prior to RNA extraction, protoscoleces
were treated with pepsin (0.5 mg/ml in Hanks' solution pH
2.0, for 3 h at 37C). cDNA was prepared by RT-PCR and
directionally cloned in a plasmid vector. RT was primed
with a tagged oligo-dT (NotI primer-adaptor). This tag was
then used to prime cDNA amplification together with a
primer for the recently characterised E. granulosus
spliced leader sequence (Brehm et al, J Biol Chem 275:
38311-8, 2000). cDNA was ligated to SalI adaptors,
digested with NotI, and cloned into pSPORT 1 cut with
SalI and NotI."

BASE COUNT 128 a 111 c 119 g 146 t
ORIGIN

Query Match 16.2%; Score 55.8; DB 13; Length 504;
Best Local Similarity 50.2%; Pred. No. 4,7e-06;
Matches 138; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 37 AGTTACGTGGAAGAATTTTCGCAAGTGCAGAGAAACCCGCGTGGAGACATATTG 96
DB 91 AATGAAGTTGAAGACACATTTAATCGATTTCAGATGATTAAGGAGATTTCAGGGCATCATTT 150
QY 97 ATCATGAATCACTCGGGTGCCTGGTGAAGAACTCGATGCTGCAGAGGGGCTTTCAG 156
DB 151 ATTATGAATTAACGATGCTGCTGCTTTCGACGACATGACAAACCAATGACATGAC 210
QY 157 TACGCTGTATTATGCAATTTGCGGAGAAAGTGCAGGCGTTCTTCCAAATGAG 216
DB 211 TATTGTGCGCTTCACAGCAACTGTTTCCAAGTCGAGGCGTGGTGTCCGTGACGGGAT 270

QY 217 CCAGCCCAAAATTTGACTCTACTGAGAGTTTCGACCAAGTATCAGAGGTGCTCATTTACA 276
DB 271 CCACGCAATGACCTCACTTTTCTGCGCGCTCCGACGCAAGATGATGATGATCGCG 330
QY 277 CCAGATGCCAAGATCAACCGTTTGGTGTTCAGAA 311
DB 331 CTTGACCAAGAGATGACACGCTGATCGTGTTCAGGA 365

Search completed: February 8, 2004, 10:21:58
Job time : 817.147 secs

THIS PAGE BLANK (USPTO)

George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G., Miranda, A., Mungall, C. J., Munoz, J., Paclet, J., Paragas, V., Park, S., Patel, S., Phouanavong, S., Pan, K., Yu, C., Lewis, S. E., Rubin, G. M., and Celisner, S.

TITLE Direct Submission
JOURNAL Submitted (22-JAN-2003) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA

COMMENT Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

FEATURES**Source**

1. 1899
/organism="Drosophila melanogaster"

/mol_type="mRNA"

/strain="y; cn bw sp"

/db_xref="taxon:7227"

1. 1899

/gene="CG8029"

/db_xref="FLYBASE:FBgn003393"

175. 1314

/gene="CG8029"

/note="Longest ORF"

/codon_start=1

/product="LD09158p"

/protein_id="AA025036.1"

/db_xref="GI:2782004"

/translation="MLMKSLIALCVIGAAVQTPVFLMGANSVAKPSLKTIVSOEFA
EQLAILEDHVVVAFRENGLSKDFICNSNOACVQALOGVSKPTIVSEVNSSEL
RSVAAKREHNSIDASKLTPPAKCAVGTALFVTRBDAESREASLESHDAIAAISQ
PECKVAALYLAAPSTAPVORRTDRDPAATGGIMKSTQFOLVYDALYNGPITV
TDKLITNSSTKLSVMDTVADKPITFDVYNGVPSLSNQLVYDNNFSSGNVAT
TFYSYCNLTLESAAVNNMYTSLFSKSLQAPDPGYKEDPFGDSMDCVGFVTPGI
LMGLFVALLLVIMFVGVCMWMDINTMDRDPDKKITTINAAAE"

BASE COUNT

446 a 522 c 443 g 488 t

ORIGIN

Query Match 99.9%; Score 1862; DB 3; Length 1899;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1862; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGCTGAGCAGCAAGCAAAATTTTCAACCAATTTTGGTTGATGTTCACTT 60
DB 1 CGAGCTGAGCAGCAAGCAAAATTTTCAACCAATTTTGGTTGATGTTCACTT 60
QY 61 CACTTTTCGTTGGTTTATGCTCTGCTGCTCTCTGCTCTCTGCTCTCTACTTAATG 120
DB 61 CACTTTTCGTTGGTTTATGCTCTCTGCTCTCTCTGCTCTCTACTTAATG 120
QY 121 TTGCTTCTCTTTTTCATTAACCCCAAGCAAGTACGATCAGCGCATGTTG 180
DB 121 TTGCTTCTCTTTTTCATTAACCCCAAGCAAGTACGATCAGCGCATGTTG 180
QY 181 TGAAGATCGCTGATGCTGTGCTCATTTGGGCTGCTGCTGCTGAGCAACGCCGCTC 240
DB 181 TGAAGATCGCTGATGCTGTGCTCATTTGGGCTGCTGCTGCTGAGCAACGCCGCTC 240

QY 241 TTTTGTGGAGCCCAAGTGTGCGAAACCTCTCTGAAGACGCTGTCCAGTTGAG 300
DB 241 TTTTGTGGAGCCCAAGTGTGCGAAACCTCTCTGAAGACGCTGTGTCCAGTTGAG 300
QY 301 TTTTCCGAGATGTTGCTGTGATTTGCTGGAAGATCAATGCTGCTGCGAATAAT 360
DB 301 TTTTCCGAGATGTTGCTGTGATTTGCTGGAAGATCAATGCTGCTGCGAATAAT 360
QY 361 GGCTGAGCAGCAAGACTTCTGTCTCCACTCCAGGCGAGTCTCTAGCCGAC 420
DB 361 GGCTGAGCAGCAAGACTTCTGTCTCCACTCCAGGCGAGTCTCTAGCCGAC 420
QY 421 CTGCAAGAGTGAAGCCCAAGACTTCTCAACCAAGCGTGAAGAACCTCTGAGCACTG 480
DB 421 CTGCAAGAGTGAAGCCCAAGACTTCTCAACCAAGCGTGAAGAACCTCTGAGCACTG 480
QY 481 CGCTCAAGTGGCCCGCAAGCGCGAGCAACTCCATGATGCTCCAGTGAAGACGACCA 540
DB 481 CGCTCAAGTGGCCCGCAAGCGCGAGCAACTCCATGATGCTCCAGTGAAGACGACCA 540
QY 541 CGGAGCAAGTGTGCGTAGGTAAGGCTGTTGTGACCTTCAAGAGCGCGCGAGAGC 600
DB 541 CGGAGCAAGTGTGCGTAGGTAAGGCTGTTGTGACCTTCAAGAGCGCGCGAGAGC 600
QY 601 CGGAGCGCAAGTGTGAGTCAATGACCGCGCATAGCTCTATCAAGCAAGCTTCAAG 660
DB 601 CGGAGCGCAAGTGTGAGTCAATGACCGCGCATAGCTCTATCAAGCAAGCTTCAAG 660
QY 661 TGAAGTGTGCTTATCTGTAACCTGCGCGCCCTCCAGCGCTCTGTGTGCGAGCGCT 720
DB 661 TGAAGTGTGCTTATCTGTAACCTGCGCGCCCTCCAGCGCTCTGTGTGCGAGCGCT 720
QY 721 ACCGCGGTGACACCGCGCGCAACCGGTGATCATGTGAGTCAAGTCAATCACTT 780
DB 721 ACCGCGGTGACACCGCGCGCAACCGGTGATCATGTGAGTCAAGTCAATCACTT 780
QY 781 CAGATCTTCAACCTGCGCTCTCAACCGGCAACCCGATCAACCTGATCACTTCAAG 840
DB 781 CAGATCTTCAACCTGCGCTCTCAACCGGCAACCCGATCAACCTGATCACTTCAAG 840
QY 841 CTGACCAACTGAGCTCTCAACAGCTTACCGCTGTTATGAGCAATCTGCGCGCAAG 900
DB 841 CTGACCAACTGAGCTCTCAACAGCTTACCGCTGTTATGAGCAATCTGCGCGCAAG 900
QY 901 CCAATCACTTTGACGTTGTTTATATGCGGATACCTTATGCTGAGCAATTTGCTTAC 960
DB 901 CCAATCACTTTGACGTTGTTTATATGCGGATACCTTATGCTGAGCAATTTGCTTAC 960
QY 961 GACAAACAACCTTCCGCTCAGCGGCGTGAATGCCCAACAGCTTCTTACTGCTGC 1020
DB 961 GACAAACAACCTTCCGCTCAGCGGCGTGAATGCCCAACAGCTTCTTACTGCTGC 1020
QY 1021 GCGAAGCTCACTCTCGAGTCCGCGCGCTCAACCAATGTAACAACCTCGAGCTTCAAG 1080
DB 1021 GCGAAGCTCACTCTCGAGTCCGCGCGCTCAACCAATGTAACAACCTCGAGCTTCAAG 1080
QY 1081 TCCCTGAGCTGAGGCTCCCTCGATGTAAGTCAAGAGGACTTTCCTTCGCGGAC 1140
DB 1081 TCCCTGAGCTGAGGCTCCCTCGATGTAAGTCAAGAGGACTTTCCTTCGCGGAC 1140
QY 1141 TCCCTGAGCTGAGGCTTCTGTAAGCGCCCGATATCTGATGGAAGCTTGTGTGCTGC 1200
DB 1141 TCCCTGAGCTGAGGCTTCTGTAAGCGCCCGATATCTGATGGAAGCTTGTGTGCTGC 1200
QY 1201 CTACTGCTGATCATGTTGCTGCGCTGTGCTGATGATGAGCATCAACAGATGAGAC 1260
DB 1201 CTACTGCTGATCATGTTGCTGCGCTGTGCTGATGATGAGCATCAACAGATGAGAC 1260
QY 1261 CGCTTGAAGATCCCAAGGCAAGACTATCAACCATCAATGCGCGCGAGTAAACCGCC 1320
DB 1261 CGCTTGAAGATCCCAAGGCAAGACTATCAACCATCAATGCGCGCGAGTAAACCGCC 1320

QY	1321	TACATCATCGGTTGACACCTTATTTGATGTTGTCAATGGCACTCAATATTCATTCCTTTA	1380
Db	1321	TACATCATCGGTTGACACCTTATTTGATGTTGTCAATGGCACTCAATATTCATTCCTTTA	1380
QY	1381	TTCCCTTCAGTCTATCCACATGTAATATGTAAAGCATCAATACCTTAATACCGTACACGAT	1440
Db	1381	TTCCCTTCAGTCTATCCACATGTAATATGTAAAGCATCAATACCTTAATACCGTACACGAT	1440
QY	1441	TACACATTAACACATACCGGGAAATTAGTCTCACAATTTGTATCGTATCCGTAAATTTTC	1500
Db	1441	TACACATTAACACATACCGGGAAATTAGTCTCACAATTTGTATCGTATCCGTAAATTTTC	1500
QY	1501	GCCGCGAAGATCAAGTTTAAGTTACGATCAACGGTTCGATTCCTGTGCGCGACAA	1560
Db	1501	GCCGCGAAGATCAAGTTTAAGTTACGATCAACGGTTCGATTCCTGTGCGCGACAA	1560
QY	1561	TTGCCTGAGTWTTCCTTGTCACACCGACGAGATCCAGATTTCAATTTAGATTAACCCCTA	1620
Db	1561	TTGCCTGAGTWTTCCTTGTCACACCGACGAGATCCAGATTTCAATTTAGATTAACCCCTA	1620
QY	1621	AATGTCGTTTATCGAAGAACCTTTAAAGACGACATGTGTGTGTGTTTAAAGCGCG	1680
Db	1621	AATGTCGTTTATCGAAGAACCTTTAAAGACGACATGTGTGTGTGTTTAAAGCGCG	1680
QY	1681	CGCGGTGTGTGTGTGCAATGCGGACGAAATCTAATCCGAAGATATCTAAGTCTTTA	1740
Db	1681	CGCGGTGTGTGTGTGCAATGCGGACGAAATCTAATCCGAAGATATCTAAGTCTTTA	1740
QY	1741	TAGTTTCTGTGTTGGCGCAATATATGCGAGACTAATTTATGCTCTGTATGATAGTTGTG	1800
Db	1741	TAGTTTCTGTGTTGGCGCAATATATGCGAGACTAATTTATGCTCTGTATGATAGTTGTG	1800
QY	1801	AATTTGTGCAACCTCTCTGTGTGTGCGACGACCAAGAAATAAAGTTTTTTTCAATGTG	1860
Db	1801	AATTTGTGCAACCTCTCTGTGTGTGCGACGACCAAGAAATAAAGTTTTTTTCAATGTG	1860
QY	1861	TT 1862	
Db	1861	TT 1862	

	AC008186	189757 bp	DNA	linear	INV 10-MAR-2001
LOCUS	AC008186				
DEFINITION	AC008186 melanogaster, chromosome 2R, region 45A-46A, BAC clone BACR0014, complete sequence.				
ACCESSION	AC008186				
VERSION	AC008186.5	GI:13270529			
KEYWORDS	HTG.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidae; Drosophilidae; Drosophila. 1 (bases 1 to 189757) Celinker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Bazzon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresner,D., Farfan,D., Ferreira,S., Frisoe,E., Gale,R.F., Garay,N.S., George,R.R., Gonzalez,M., Houch,J., Hoskins,K.A., Hostin,D., Howland,T.U., Iadegou,C., Jalali,M., Krusey,D., Li,P., Mactel,B., Koshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nuno,U., Pated,J.V., Paragás,V., Park,S., Patel,S., Pfeiffer,B., Phonathanavong,S., Plittman,G.S., Putrì,V., Richards,S., Scheeler,F., Stjepanovic,M., Strong,R., Swirkkaas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.				
TITLE	Sequencing of Drosophila chromosome 2R, region 45A-46A				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 189757)				
AUTHORS	Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazer,J.G., Butenhoff,C., Champe,M., Chavez,C., Chew,W., Ciesiolka,L.,				

Doyle, C. M., Fattan, D. E., Galle, R., George, R. A., Harris, N. L., Hoskins, R. A., Houston, C. A., Hummasti, S. R., Karra, K., Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomutan, M. A., Mazza, P., Moshrefi, A. R., Moshrefi, M., Nixon, K., Pacled, J. M., Park, S., Pfeiffer, B., Poon, L., Segueira, A., Sethi, H., Snij, E., Stierikas, R. R., Wan, K. H., Weinburg, T., Zhang, R., Zlerian, L. L. and Rubin, G. M.

TITLE	Direct Submission
JOURNAL	Submitted (29-UTL-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT	On Mar 10, 2001 this sequence version replaced gi:6997001.

Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-1211
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.frititfly.org/sequence/>) or send email to bdg@frititfly.berkeley.edu.

FEATURES	Location/Qualifiers
source	1. .189757

```
/organism="Drosophila melanogaster"  
/mol_type="genomic DNA"  
/strain="y; cn bw sp"  
/db_xref="taxon:7227"
```

```

/chromosome="2R"
/map="45A-46A"
/clone="BACR10M14 (D583)"
/clone_1b="PPCt-98 (Roswell) Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6"

```

BASE COUNT	52250	a	42350	c	42896	g	52261	t
ORIGIN								

Query Match	87.6%	Score 1632	DB 3	Length 189757
-------------	-------	------------	------	---------------

Best Local Similarity 90.3%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 0; Indels 201; Gaps 3,

1 CGACGTGAGCAAGGCAGAAATTTTACACCAATTTTGTGTGATTCAGTT 60

Db 135296 CGACGTGAGCAAGCGAAATTTTCACACCATTTTCGTTGTAGTTCACTT 135

61 CACTTTCGGTGGTTTTAGTGCCTCTGCGCTCTCTGGTTCCTTAATGCAATT 120

Db 13536 CACITTTGGITGTTTTAGTGCCTCTTGGCTTCTCCITTGTTCCITACITATGCAATT 1354

121 TTGGTCTCCTTGTGTTCCATTAAACCCACCGAAGTAAGCGATTCAGGCGCATGTTG 180

D6 135416 TTTCGTTCTCCCTTGTTCATTTAAACCCACCGAAGTAAGCAATCCAGCGCATGTTC 135

181 TGGAAATTCGCTGATTCGGTGTGCGGACATGGGCGCGCCGCGGAGGCAAAAGCCCGCC 240

00	241	TTTTTTTGGGGGAGCCCA	256
01		-----	
02		-----	
03		-----	
04		-----	
05		-----	
06		-----	
07		-----	
08		-----	
09		-----	
10		-----	
11		-----	
12		-----	
13		-----	
14		-----	
15		-----	
16		-----	
17		-----	
18		-----	
19		-----	
20		-----	
21		-----	
22		-----	
23		-----	
24		-----	
25		-----	
26		-----	
27		-----	
28		-----	
29		-----	
30		-----	
31		-----	
32		-----	
33		-----	
34		-----	
35		-----	
36		-----	
37		-----	
38		-----	
39		-----	
40		-----	
41		-----	
42		-----	
43		-----	
44		-----	
45		-----	
46		-----	
47		-----	
48		-----	
49		-----	
50		-----	
51		-----	
52		-----	
53		-----	
54		-----	
55		-----	
56		-----	
57		-----	
58		-----	
59		-----	
60		-----	
61		-----	
62		-----	
63		-----	
64		-----	
65		-----	
66		-----	
67		-----	
68		-----	
69		-----	
70		-----	
71		-----	
72		-----	
73		-----	
74		-----	
75		-----	
76		-----	
77		-----	
78		-----	
79		-----	
80		-----	
81		-----	
82		-----	
83		-----	
84		-----	
85		-----	
86		-----	
87		-----	
88		-----	
89		-----	
90		-----	
91		-----	
92		-----	
93		-----	
94		-----	
95		-----	
96		-----	
97		-----	
98		-----	
99		-----	
100		-----	

Db 135536 TTTTGTGGGACCAA CAGTGGGTGACGCGCTGCACCACTTGAATGTGAGTGTACGA 135537

257 -----ACAGTGTGCGAAACCTCCCTGAAGACGGTGTCCCAAGT 297

D5
135596 GTGCTTCCTCTTTTTCACAGTGTGGCGAACCCTCCTGAAGACGGTGTCCCAAGTG 1356

298 GAGTTTGCAGACAGTTGGCTGCATTGCTGGAAGATCAATGTGTCGTGGCCTTCGAGGAA 357

Db 135656 GAGTTGCCGAGAGTTGGCTGCATTCCTGGAGATCATGTGTCGTGACCTTCGAGGA 135

QY 358 AATGC----- 363
|||||

D6 135716 AATGGCCGTAAGTCTTGAGAACCCACTTAGATAGCTAGGGCTTGACACATGTGTTCG 1357

364 -----CTGAGCAGCAGAGACTTCCTGCTCCTCAACTCCAGGCGAGTCTGC 411

Db	135776	TCCCCACTCAGCTGAGCAGCAGGACTTTCCTGTGCTCCAACTCCAGGCGCAGTCTCG	135835
Qy	412	TACGCCCAGCTGCGAGGAGTGAAGCCCCAAAGACTTACATCACAGCGTGGAGAACCCCTCG	471
Db	135836	TACGCCCAGCTGCGAGGAGTGAAGCCCCAAAGACTTACATCACAGCGTGGAGAACCCCTCG	135895
Qy	472	GAGGCACTGCGCTCAGTGTGGCCCGCCAAAGCGCGAGCACAATCCATCGATGCGATGGAAG	531
Db	135896	GAGGCACTGCGCTCAGTGTGGCCCGCCAAAGCGCGAGCACAATCCATCGATGCGATGGAAG	135955
Qy	532	CTGACACAGCGCGGCCAAGTGTGCCGATGAGGAGGCGCCCTGTTTGTGACTTTCGAGAGCGCC	591
Db	135956	CTGACACAGCGCGGCCAAGTGTGCCGATGAGGAGGCGCCCTGTTTGTGACTTTCGAGAGCGCC	136015
Qy	592	GCCGAGAGCCGGAGGCGCAGTCTGAGTCAAT-----	624
Db	136016	GCCGAGAGCCGGAGGCGCAGTCTGAGTCAATGATGGGTCAACAATTAAGTACTACAA	136075
Qy	625	-----GACGCCGCATAGCT	639
Db	136076	GCGCTCAATTAAGACACACTAATTTAACTGATTTAAACAATTAAACCCCGCATAGCT	136135
Qy	640	GCTATCAGCAACAGTTGAGTGAAGGATGCAAGTGGCTTATCTGTACCTGGCCGCCCTCCACG	699
Db	136136	GCTATCAGCAACAGTTGAGTGAAGGATGCAAGTGGCTTATCTGTACCTGGCCGCCCTCCACG	136195
Qy	700	GCTCCGTGTGATCAGCGTGTGATCCGCGCTGACACCGCGGCACAAACCGGTGGATCATG	759
Db	136196	GCTCCGTGTGATCAGCGTGTGATCCGCGCTGACACCGCGGCACAAACCGGTGGATCATG	136255
Qy	760	TGGAAGTGCACCAATCATGTTTCAGATCTTCTACATGCTCCCTGCTTACACGCGCAACCC	819
Db	136256	TGGAAGTGCACCAATCATGTTTCAGATCTTCTACATGCTCCCTGCTTACACGCGCAACCC	136315
Qy	820	ATCACCCTGCACCGACTCAGCTCAACCACTAGTCTACCAAGCTATCCGTGTATG	879
Db	136316	ATCACCCTGCACCGACTCAGCTCAGCTCAACCACTAGTCTACCAAGCTATCCGTGTATG	136375
Qy	880	GACACATCTGTGCGCGACAAGCCAACTCACCTTGTGACGTTGTATATAGCGGATACCTT	939
Db	136376	GACACATCTGTGCGCGACAAGCCAACTCACCTTGTGACGTTGTATATAGCGGATACCTT	136435
Qy	940	AGCTGAGCAATTTGGTCTACGACAAACAACCTTCGCTCAGCGCGGTGAATGCCCA	999
Db	136436	AGCTGAGCAATTTGGTCTACGACAAACAACCTTCGCTCAGCGCGGTGAATGCCCA	136495
Qy	1000	ACCAGTTCTCTTACTCGTGGCGCAACCTACCTCGAGTGGCGGCGCGTCAACAATG	1059
Db	136496	ACCAGTTCTCTTACTCGTGGCGCAACCTACCTCGAGTGGCGGCGCGTCAACAATG	136555
Qy	1060	TACACAACCCCTGAGCTTCAAGTCCCTGAGAGCTGCAAGGCTCCCTGCATGTGACATCAAG	1119
Db	136556	TACACAACCCCTGAGCTTCAAGTCCCTGAGAGCTGCAAGGCTCCCTGCATGTGACATCAAG	136615
Qy	1120	GAGGACTTTCCTTCGCGCACTCTCGGACGTGTGTGGCTTCTGTGACGCCCGGATCTTG	1179
Db	136616	GAGGACTTTCCTTCGCGCACTCTCGGACGTGTGTGGCTTCTGTGACGCCCGGATCTTG	136675
Qy	1180	ATGGGACTGTTTGTGTGCTGCGCCCTACGTGCTGTATCATGTTGCTGGCGTCTGTGATG	1239
Db	136676	ATGGGACTGTTTGTGTGCTGCGCCCTACGTGCTGTATCATGTTGCTGGCGTCTGTGATG	136735
Qy	1240	ATGGGACATCAACGATGAGCGGCTTGCAGATCCCAAGGCAAGACTACACATCAAT	1299
Db	136736	ATGGGACATCAACGATGAGCGGCTTGCAGATCCCAAGGCAAGACTACACATCAAT	136795
Qy	1300	GCGCGCGCAGATGAACCGGCTTACATCATCGTTCACACTTATTTGATGTTCAATGCG	1359
Db	136796	GCGCGCGCAGATGAACCGGCTTACATCATCGTTCACACTTATTTGATGTTCAATGCG	136855
Qy	1360	CACCTAATATTCATCTCTTATTCCTTCACTATCAATGTATATGTTAAAGATCAA	1419

Db	136856	CAGTCMAATTCATTCCTTTATTCCTTGAGTCCTATCCAGATGTATATATGTTAAAGCATCA	136915
Qy	1420	TACTTAATACCGGTACAGCATTTACACATPAACATPAACCGGGAAATTAGTCTCACAAT	1479
Db	136916	TACTTAATACCGGTACAGCATTTACACATPAACATPAACCGGGAAATTAGTCTCACAAT	136975
Qy	1480	TGTATCGATCCGTAAATTTTGGCGCCCAAGATCAGTTTACCGTCAGCATCAAGCGTGG	1539
Db	136976	TGTATCGATCCGTAAATTTTGGCGCCCAAGATCAGTTTACCGTCAGCATCAAGCGTGG	137035
Qy	1540	CTGATTCCTTGTGCGCGACAAATTTGCTTGAAGTTTGTGTCACCGACGACATCCAGTA	1599
Db	137036	CTGATTCCTTGTGCGCGACAAATTTGCTTGAAGTTTGTGTCACCGACGACATCCAGTA	137095
Qy	1600	TTTCATTTAGATATAGCCCTTAATGTGCTTATGCAAGAACTTTTAAAGAAGCATATGT	1659
Db	137096	TTTCATTTAGATATAGCCCTTAATGTGCTTATGCAAGAACTTTTAAAGAAGCATATGT	137155
Qy	1660	GTGTGTGTGTTTAAAGAGCGCGCGCTGTGTGTGTGTGTGTGCACATGGCAGTAAATCTTAATCC	1719
Db	137156	GTGTGTGTGTTTAAAGAGCGCGCGCTGTGTGTGTGTGTGTGCACATGGCAGTAAATCTTAATCC	137215
Qy	1720	GAAGTATATCTTAAGCTTTTAAAGTTTCTGTGTGGCGCAATATATGCGAGACTAATTTAT	1779
Db	137216	GAAGTATATCTTAAGCTTTTAAAGTTTCTGTGTGGCGCAATATATGCGAGACTAATTTAT	137275
Qy	1780	GGCTCTGATGATAGTTTGTGCTATTTGTGTGCACCTCTGTGTGTGTGTGTGCAGGCCCAAGAAA	1839
Db	137276	GGCTCTGATGATAGTTTGTGCTATTTGTGTGCACCTCTGTGTGTGTGTGTGCAGGCCCAAGAAA	137335
Qy	1840	ATTAAGTTTTTTTCAATGTGTGTTA	1863
Db	137336	ATTAAGTTTTTTTCAATGTGTGTTA	137359

RESULT 3	24313 bp	DNA	linear	INV 14-FEB-2003
LOCUS	AE003834			
DEFINITION	Drosophila melanogaster chromosome 2R, section 16 of 74 of the complete sequence.			
ACCESSION	AE003834	AE002787	AE013599	
VERSION	AE003834.3	GI:21627610		
KEYWORDS				
SOURCE				
ORGANISM	Drosophila melanogaster (fruit fly)			
REFERENCE	Drosophila melanogaster			
AUTHORS	<p>Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.</p> <p>1 (bases 1 to 243313)</p> <p>Adams M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Wortman, J.R., Yandell, M.D., Zhang, Q., Chen, L.X., Brandon, R.C., Rogers, Y.H., Blazer, R.G., Champe, M., Pfeiffer, B.D., Mani, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Gabor, G.L., Abail, J.F., Agbayani, A., An, H.J., Andrews, Pflamknoch, C., Baldwin, D., Ballew, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y., Benos, P.V., Bernan, B.P., Bhandari, D., Bolshakov, S., Borkov, D., Botchan, M.R., Bouck, J., Brooksstein, P., Brotlier, P., Burris, K.C., Busam, D.A., Butler, H., Cantlev, E., Center, A., Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de Pablo, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M., Dodson, K., Doup, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferreira, S., Fleischmann, W., Foster, C., Gabrielian, A.E., Garg, N.S., Gelbar, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Honck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G.H., Ke, Z., Kemison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Laslo, P., Lei, Y., Levitsky, A.A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Liu, X., Mattei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Mekulov, G., Milshina, N.V., Mobarry, C.,</p>			

Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusser, D.R., Pacle, J.M., Palazzolo, M., Pittman, G.S., Pan, S., Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stappleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Waasman, D.A., Weinstein, G.M., Weissenbach, J., Williams, S.M., Woodruff, W., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, X., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
The genome sequence of *Drosophila melanogaster*
Science 287 (5461), 2185-2195 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
20196006
10731132

REFERENCE
AUTHORS
2 (bases 1 to 243313)
Celniker, S.E., Adams, M.D., Kronmiller, B., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., Banzon, J., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorset, V., Doup, L.E., Doyle, C., Dreonek, D., Fatfan, D., Ferreira, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houch, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jaitai, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nuno, J., Pacle, J., Parag, V., Park, S., Patel, S., Pfeiffer, B., Phouenavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Shapiro, M., Strong, R., Svirskas, R., Tector, C., Tyler, D., Williams, S.M., Zaveri, J.S., Smith, H.O., Venter, J.C. and Rubin, G.M.
Sequencing of *Drosophila melanogaster* genome
Unpublished

TITLE
JOURNAL
AUTHORS
3 (bases 1 to 243313)
Mist, S., Crosby, M.A., Matthews, B.B., Bayraktaroglu, L., Campbell, K., Hradecky, P., Huang, Y., Kaminker, J.S., Prochuk, S.E., Smith, C.D., Tupy, J.L., Bergman, C.M., Berman, B.P., Carlson, J.W., Celniker, S.E., Clamp, M.E., Drysdale, R.A., Emmert, D., Frise, E., de Grey, A.D.N.J., Harris, N.L., Kronmiller, B., Marshall, B., Millburn, G.H., Richter, J., Russo, S., Searle, S.M.J., Smith, E., Shu, S., Smutnick, F., Whitfield, E.J., Ashburner, M., Gelbart, W.M., Rubin, G.M., Munhall, C.J. and Lewis, S.E.
Annotation of *Drosophila melanogaster* genome
Unpublished

TITLE
JOURNAL
AUTHORS
4 (bases 1 to 243313)
Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, J.C.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

REFERENCE
JOURNAL
5 (bases 1 to 243313)
FlyBase
Direct Submission
Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA

TITLE
JOURNAL
6 (bases 1 to 243313)
FlyBase
Direct Submission
Submitted (23-JAN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA

COMMENT
FEATURES
On Jun 28, 2002 this sequence version replaced gi:10727693.
location/Qualifiers
1..243313
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="2R"
/note="genotype: y[1]; cn[1] bw[1] sp[1]; Rh6[1]"
3648..12101
/gene="ana"
/locus_tag="CG8084"
/note="last curated on Wed Mar 13 09:30:29 PST 2002"

gene

map="45A1-45A1"
/db_xref="FLYBASE:FBgn0011746"
join(3648..4173,8618..8997,9387..9635,10420..10492,10917..12101)
/gene="ana"
/locus_tag="CG8084"
/product="CG8084-Ra"
/db_xref="FLYBASE:FBgn0011746"
join(13873..4173,8618..8997,9387..9635,10420..10492,10917..11343)
/gene="ana"
/locus_tag="CG8084"
/note="ana gene product"
/codon_start=1
/product="CG8084-PA"
/db_xref="GI:7303954"
/db_xref="FLYBASE:FBgn0011746"
/translation="MASAMRGKCRSRIRLVLILSLITMGADSRAPFPDPSFIEGVOSEVNPFRITLIRFNLEKQILISIQNRSNPMDDSAQSNQOYLQAVATORDIFRVOKALSNENPGASKEKAGFPICNAETTPEDMSIGNVTLOFASVPSINDDRLSSALIRIKTNPGOTREHNPQASTOPISTENGNTA PNCADPVPQIPQIVYSI VHQKKQKRTCTNTAMSSSTGVEIDYKCALAYEQQRQOQPOQOLTA SVGLIMEVHDEENLIRPGIFAPPTCDQADIAVPSVYTEPKSKLAWTLPRK PLDIFPNSNMKNSYNTPKSRAPLESTTSPITNDQSDSEYSSQQRVAPLVR HRRHHNHQCPSESESAQLEAEIEAEIEMSSASNSQOQMEPISNHRHRTGHHPHH OLOHHHHHRTKHKRIIPAKHQE"
13293..16922
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="last curated on Mon Feb 11 10:24:48 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0025709"
join(13293..13488,14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/product="CG8083-Ra"
/db_xref="FLYBASE:FBgn0025709"
join(14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="BEST:CK00459 gene product"
/codon_start=1
/product="CG8083-PA"
/protein_id="AAP58997.1"
/db_xref="GI:7303953"
/db_xref="FLYBASE:FBgn0025709"
/translation="MSAESKGAINNGLDHEKELDIRDSEETHEPLDTSVANKKGYEPKPKVRLVRSISIVLHLGVGYSVATVHYHDTNTECKMKNPLGGINCTGYMLLLIGFIYLGFIYTVYFKPMVGHSHLNRYIKPSKKHNSRTRVSLASIAL LIALALIVYFECRDQTLVSLVAPCEPILCGYVSTNRRALNMIVITGTICQDL GFCIRMEGRKIFBCLGNKRVATFLGATGDFVGNVNVFAFALIVIFPFS FPISTLIYMGVQWVILKGMILIOQILGTVCESVYAAANIFLGMSESLIRPYINK LKSRHSHVMSGPAVSGTVAAVLSFGASAAHLITSVMAAPATLAIISKYMPETE BSOTSDSTEELKQSDSLDAASGASNAVIVIGIITVAFAVFAFLNGVSWF GTVGLGQIDFEWIFSKFLPLVPMAGCPKEDCDITAKVATKTIINNEVAYERLQOI IENNDITARSAGIATPAICGFANPSSLGILISLSAMAPHRSTTVAFAFVAGSI VCFVASFGAGILIOBDERANYNRIPEKIGRKNVQF"
16984..18940
/locus_tag="CG11778"
/note="last curated on Wed Mar 13 10:09:14 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0033371"
join(16984..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778-Ra"
/product="CG11778-PA"
/db_xref="FLYBASE:FBgn0033371"
join(17021..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778"
/note="CG11778 gene product"
/codon_start=1

gene

map="45A1-45A1"
/db_xref="FLYBASE:FBgn0011746"
join(3648..4173,8618..8997,9387..9635,10420..10492,10917..12101)
/gene="ana"
/locus_tag="CG8084"
/product="CG8084-Ra"
/db_xref="FLYBASE:FBgn0011746"
join(13873..4173,8618..8997,9387..9635,10420..10492,10917..11343)
/gene="ana"
/locus_tag="CG8084"
/note="ana gene product"
/codon_start=1
/product="CG8084-PA"
/db_xref="GI:7303954"
/db_xref="FLYBASE:FBgn0011746"
/translation="MASAMRGKCRSRIRLVLILSLITMGADSRAPFPDPSFIEGVOSEVNPFRITLIRFNLEKQILISIQNRSNPMDDSAQSNQOYLQAVATORDIFRVOKALSNENPGASKEKAGFPICNAETTPEDMSIGNVTLOFASVPSINDDRLSSALIRIKTNPGOTREHNPQASTOPISTENGNTA PNCADPVPQIPQIVYSI VHQKKQKRTCTNTAMSSSTGVEIDYKCALAYEQQRQOQPOQOLTA SVGLIMEVHDEENLIRPGIFAPPTCDQADIAVPSVYTEPKSKLAWTLPRK PLDIFPNSNMKNSYNTPKSRAPLESTTSPITNDQSDSEYSSQQRVAPLVR HRRHHNHQCPSESESAQLEAEIEAEIEMSSASNSQOQMEPISNHRHRTGHHPHH OLOHHHHHRTKHKRIIPAKHQE"
13293..16922
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="last curated on Mon Feb 11 10:24:48 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0025709"
join(13293..13488,14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/product="CG8083-Ra"
/db_xref="FLYBASE:FBgn0025709"
join(14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="BEST:CK00459 gene product"
/codon_start=1
/product="CG8083-PA"
/protein_id="AAP58997.1"
/db_xref="GI:7303953"
/db_xref="FLYBASE:FBgn0025709"
/translation="MSAESKGAINNGLDHEKELDIRDSEETHEPLDTSVANKKGYEPKPKVRLVRSISIVLHLGVGYSVATVHYHDTNTECKMKNPLGGINCTGYMLLLIGFIYLGFIYTVYFKPMVGHSHLNRYIKPSKKHNSRTRVSLASIAL LIALALIVYFECRDQTLVSLVAPCEPILCGYVSTNRRALNMIVITGTICQDL GFCIRMEGRKIFBCLGNKRVATFLGATGDFVGNVNVFAFALIVIFPFS FPISTLIYMGVQWVILKGMILIOQILGTVCESVYAAANIFLGMSESLIRPYINK LKSRHSHVMSGPAVSGTVAAVLSFGASAAHLITSVMAAPATLAIISKYMPETE BSOTSDSTEELKQSDSLDAASGASNAVIVIGIITVAFAVFAFLNGVSWF GTVGLGQIDFEWIFSKFLPLVPMAGCPKEDCDITAKVATKTIINNEVAYERLQOI IENNDITARSAGIATPAICGFANPSSLGILISLSAMAPHRSTTVAFAFVAGSI VCFVASFGAGILIOBDERANYNRIPEKIGRKNVQF"
16984..18940
/locus_tag="CG11778"
/note="last curated on Wed Mar 13 10:09:14 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0033371"
join(16984..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778-Ra"
/product="CG11778-PA"
/db_xref="FLYBASE:FBgn0033371"
join(17021..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778"
/note="CG11778 gene product"
/codon_start=1

gene

map="45A1-45A1"
/db_xref="FLYBASE:FBgn0011746"
join(3648..4173,8618..8997,9387..9635,10420..10492,10917..12101)
/gene="ana"
/locus_tag="CG8084"
/product="CG8084-Ra"
/db_xref="FLYBASE:FBgn0011746"
join(13873..4173,8618..8997,9387..9635,10420..10492,10917..11343)
/gene="ana"
/locus_tag="CG8084"
/note="ana gene product"
/codon_start=1
/product="CG8084-PA"
/db_xref="GI:7303954"
/db_xref="FLYBASE:FBgn0011746"
/translation="MASAMRGKCRSRIRLVLILSLITMGADSRAPFPDPSFIEGVOSEVNPFRITLIRFNLEKQILISIQNRSNPMDDSAQSNQOYLQAVATORDIFRVOKALSNENPGASKEKAGFPICNAETTPEDMSIGNVTLOFASVPSINDDRLSSALIRIKTNPGOTREHNPQASTOPISTENGNTA PNCADPVPQIPQIVYSI VHQKKQKRTCTNTAMSSSTGVEIDYKCALAYEQQRQOQPOQOLTA SVGLIMEVHDEENLIRPGIFAPPTCDQADIAVPSVYTEPKSKLAWTLPRK PLDIFPNSNMKNSYNTPKSRAPLESTTSPITNDQSDSEYSSQQRVAPLVR HRRHHNHQCPSESESAQLEAEIEAEIEMSSASNSQOQMEPISNHRHRTGHHPHH OLOHHHHHRTKHKRIIPAKHQE"
13293..16922
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="last curated on Mon Feb 11 10:24:48 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0025709"
join(13293..13488,14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/product="CG8083-Ra"
/db_xref="FLYBASE:FBgn0025709"
join(14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="BEST:CK00459 gene product"
/codon_start=1
/product="CG8083-PA"
/protein_id="AAP58997.1"
/db_xref="GI:7303953"
/db_xref="FLYBASE:FBgn0025709"
/translation="MSAESKGAINNGLDHEKELDIRDSEETHEPLDTSVANKKGYEPKPKVRLVRSISIVLHLGVGYSVATVHYHDTNTECKMKNPLGGINCTGYMLLLIGFIYLGFIYTVYFKPMVGHSHLNRYIKPSKKHNSRTRVSLASIAL LIALALIVYFECRDQTLVSLVAPCEPILCGYVSTNRRALNMIVITGTICQDL GFCIRMEGRKIFBCLGNKRVATFLGATGDFVGNVNVFAFALIVIFPFS FPISTLIYMGVQWVILKGMILIOQILGTVCESVYAAANIFLGMSESLIRPYINK LKSRHSHVMSGPAVSGTVAAVLSFGASAAHLITSVMAAPATLAIISKYMPETE BSOTSDSTEELKQSDSLDAASGASNAVIVIGIITVAFAVFAFLNGVSWF GTVGLGQIDFEWIFSKFLPLVPMAGCPKEDCDITAKVATKTIINNEVAYERLQOI IENNDITARSAGIATPAICGFANPSSLGILISLSAMAPHRSTTVAFAFVAGSI VCFVASFGAGILIOBDERANYNRIPEKIGRKNVQF"
16984..18940
/locus_tag="CG11778"
/note="last curated on Wed Mar 13 10:09:14 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0033371"
join(16984..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778-Ra"
/product="CG11778-PA"
/db_xref="FLYBASE:FBgn0033371"
join(17021..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778"
/note="CG11778 gene product"
/codon_start=1

gene

map="45A1-45A1"
/db_xref="FLYBASE:FBgn0011746"
join(3648..4173,8618..8997,9387..9635,10420..10492,10917..12101)
/gene="ana"
/locus_tag="CG8084"
/product="CG8084-Ra"
/db_xref="FLYBASE:FBgn0011746"
join(13873..4173,8618..8997,9387..9635,10420..10492,10917..11343)
/gene="ana"
/locus_tag="CG8084"
/note="ana gene product"
/codon_start=1
/product="CG8084-PA"
/db_xref="GI:7303954"
/db_xref="FLYBASE:FBgn0011746"
/translation="MASAMRGKCRSRIRLVLILSLITMGADSRAPFPDPSFIEGVOSEVNPFRITLIRFNLEKQILISIQNRSNPMDDSAQSNQOYLQAVATORDIFRVOKALSNENPGASKEKAGFPICNAETTPEDMSIGNVTLOFASVPSINDDRLSSALIRIKTNPGOTREHNPQASTOPISTENGNTA PNCADPVPQIPQIVYSI VHQKKQKRTCTNTAMSSSTGVEIDYKCALAYEQQRQOQPOQOLTA SVGLIMEVHDEENLIRPGIFAPPTCDQADIAVPSVYTEPKSKLAWTLPRK PLDIFPNSNMKNSYNTPKSRAPLESTTSPITNDQSDSEYSSQQRVAPLVR HRRHHNHQCPSESESAQLEAEIEAEIEMSSASNSQOQMEPISNHRHRTGHHPHH OLOHHHHHRTKHKRIIPAKHQE"
13293..16922
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="last curated on Mon Feb 11 10:24:48 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0025709"
join(13293..13488,14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/product="CG8083-Ra"
/db_xref="FLYBASE:FBgn0025709"
join(14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="BEST:CK00459 gene product"
/codon_start=1
/product="CG8083-PA"
/protein_id="AAP58997.1"
/db_xref="GI:7303953"
/db_xref="FLYBASE:FBgn0025709"
/translation="MSAESKGAINNGLDHEKELDIRDSEETHEPLDTSVANKKGYEPKPKVRLVRSISIVLHLGVGYSVATVHYHDTNTECKMKNPLGGINCTGYMLLLIGFIYLGFIYTVYFKPMVGHSHLNRYIKPSKKHNSRTRVSLASIAL LIALALIVYFECRDQTLVSLVAPCEPILCGYVSTNRRALNMIVITGTICQDL GFCIRMEGRKIFBCLGNKRVATFLGATGDFVGNVNVFAFALIVIFPFS FPISTLIYMGVQWVILKGMILIOQILGTVCESVYAAANIFLGMSESLIRPYINK LKSRHSHVMSGPAVSGTVAAVLSFGASAAHLITSVMAAPATLAIISKYMPETE BSOTSDSTEELKQSDSLDAASGASNAVIVIGIITVAFAVFAFLNGVSWF GTVGLGQIDFEWIFSKFLPLVPMAGCPKEDCDITAKVATKTIINNEVAYERLQOI IENNDITARSAGIATPAICGFANPSSLGILISLSAMAPHRSTTVAFAFVAGSI VCFVASFGAGILIOBDERANYNRIPEKIGRKNVQF"
16984..18940
/locus_tag="CG11778"
/note="last curated on Wed Mar 13 10:09:14 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0033371"
join(16984..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778-Ra"
/product="CG11778-PA"
/db_xref="FLYBASE:FBgn0033371"
join(17021..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778"
/note="CG11778 gene product"
/codon_start=1

gene

map="45A1-45A1"
/db_xref="FLYBASE:FBgn0011746"
join(3648..4173,8618..8997,9387..9635,10420..10492,10917..12101)
/gene="ana"
/locus_tag="CG8084"
/product="CG8084-Ra"
/db_xref="FLYBASE:FBgn0011746"
join(13873..4173,8618..8997,9387..9635,10420..10492,10917..11343)
/gene="ana"
/locus_tag="CG8084"
/note="ana gene product"
/codon_start=1
/product="CG8084-PA"
/db_xref="GI:7303954"
/db_xref="FLYBASE:FBgn0011746"
/translation="MASAMRGKCRSRIRLVLILSLITMGADSRAPFPDPSFIEGVOSEVNPFRITLIRFNLEKQILISIQNRSNPMDDSAQSNQOYLQAVATORDIFRVOKALSNENPGASKEKAGFPICNAETTPEDMSIGNVTLOFASVPSINDDRLSSALIRIKTNPGOTREHNPQASTOPISTENGNTA PNCADPVPQIPQIVYSI VHQKKQKRTCTNTAMSSSTGVEIDYKCALAYEQQRQOQPOQOLTA SVGLIMEVHDEENLIRPGIFAPPTCDQADIAVPSVYTEPKSKLAWTLPRK PLDIFPNSNMKNSYNTPKSRAPLESTTSPITNDQSDSEYSSQQRVAPLVR HRRHHNHQCPSESESAQLEAEIEAEIEMSSASNSQOQMEPISNHRHRTGHHPHH OLOHHHHHRTKHKRIIPAKHQE"
13293..16922
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="last curated on Mon Feb 11 10:24:48 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0025709"
join(13293..13488,14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/product="CG8083-Ra"
/db_xref="FLYBASE:FBgn0025709"
join(14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="BEST:CK00459 gene product"
/codon_start=1
/product="CG8083-PA"
/protein_id="AAP58997.1"
/db_xref="GI:7303953"
/db_xref="FLYBASE:FBgn0025709"
/translation="MSAESKGAINNGLDHEKELDIRDSEETHEPLDTSVANKKGYEPKPKVRLVRSISIVLHLGVGYSVATVHYHDTNTECKMKNPLGGINCTGYMLLLIGFIYLGFIYTVYFKPMVGHSHLNRYIKPSKKHNSRTRVSLASIAL LIALALIVYFECRDQTLVSLVAPCEPILCGYVSTNRRALNMIVITGTICQDL GFCIRMEGRKIFBCLGNKRVATFLGATGDFVGNVNVFAFALIVIFPFS FPISTLIYMGVQWVILKGMILIOQILGTVCESVYAAANIFLGMSESLIRPYINK LKSRHSHVMSGPAVSGTVAAVLSFGASAAHLITSVMAAPATLAIISKYMPETE BSOTSDSTEELKQSDSLDAASGASNAVIVIGIITVAFAVFAFLNGVSWF GTVGLGQIDFEWIFSKFLPLVPMAGCPKEDCDITAKVATKTIINNEVAYERLQOI IENNDITARSAGIATPAICGFANPSSLGILISLSAMAPHRSTTVAFAFVAGSI VCFVASFGAGILIOBDERANYNRIPEKIGRKNVQF"
16984..18940
/locus_tag="CG11778"
/note="last curated on Wed Mar 13 10:09:14 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0033371"
join(16984..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778-Ra"
/product="CG11778-PA"
/db_xref="FLYBASE:FBgn0033371"
join(17021..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778"
/note="CG11778 gene product"
/codon_start=1

gene

map="45A1-45A1"
/db_xref="FLYBASE:FBgn0011746"
join(3648..4173,8618..8997,9387..9635,10420..10492,10917..12101)
/gene="ana"
/locus_tag="CG8084"
/product="CG8084-Ra"
/db_xref="FLYBASE:FBgn0011746"
join(13873..4173,8618..8997,9387..9635,10420..10492,10917..11343)
/gene="ana"
/locus_tag="CG8084"
/note="ana gene product"
/codon_start=1
/product="CG8084-PA"
/db_xref="GI:7303954"
/db_xref="FLYBASE:FBgn0011746"
/translation="MASAMRGKCRSRIRLVLILSLITMGADSRAPFPDPSFIEGVOSEVNPFRITLIRFNLEKQILISIQNRSNPMDDSAQSNQOYLQAVATORDIFRVOKALSNENPGASKEKAGFPICNAETTPEDMSIGNVTLOFASVPSINDDRLSSALIRIKTNPGOTREHNPQASTOPISTENGNTA PNCADPVPQIPQIVYSI VHQKKQKRTCTNTAMSSSTGVEIDYKCALAYEQQRQOQPOQOLTA SVGLIMEVHDEENLIRPGIFAPPTCDQADIAVPSVYTEPKSKLAWTLPRK PLDIFPNSNMKNSYNTPKSRAPLESTTSPITNDQSDSEYSSQQRVAPLVR HRRHHNHQCPSESESAQLEAEIEAEIEMSSASNSQOQMEPISNHRHRTGHHPHH OLOHHHHHRTKHKRIIPAKHQE"
13293..16922
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="last curated on Mon Feb 11 10:24:48 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0025709"
join(13293..13488,14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/product="CG8083-Ra"
/db_xref="FLYBASE:FBgn0025709"
join(14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="BEST:CK00459 gene product"
/codon_start=1
/product="CG8083-PA"
/protein_id="AAP58997.1"
/db_xref="GI:7303953"
/db_xref="FLYBASE:FBgn0025709"
/translation="MSAESKGAINNGLDHEKELDIRDSEETHEPLDTSVANKKGYEPKPKVRLVRSISIVLHLGVGYSVATVHYHDTNTECKMKNPLGGINCTGYMLLLIGFIYLGFIYTVYFKPMVGHSHLNRYIKPSKKHNSRTRVSLASIAL LIALALIVYFECRDQTLVSLVAPCEPILCGYVSTNRRALNMIVITGTICQDL GFCIRMEGRKIFBCLGNKRVATFLGATGDFVGNVNVFAFALIVIFPFS FPISTLIYMGVQWVILKGMILIOQILGTVCESVYAAANIFLGMSESLIRPYINK LKSRHSHVMSGPAVSGTVAAVLSFGASAAHLITSVMAAPATLAIISKYMPETE BSOTSDSTEELKQSDSLDAASGASNAVIVIGIITVAFAVFAFLNGVSWF GTVGLGQIDFEWIFSKFLPLVPMAGCPKEDCDITAKVATKTIINNEVAYERLQOI IENNDITARSAGIATPAICGFANPSSLGILISLSAMAPHRSTTVAFAFVAGSI VCFVASFGAGILIOBDERANYNRIPEKIGRKNVQF"
16984..18940
/locus_tag="CG11778"
/note="last curated on Wed Mar 13 10:09:14 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0033371"
join(16984..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778-Ra"
/product="CG11778-PA"
/db_xref="FLYBASE:FBgn0033371"
join(17021..17165,17217..17532,17759..18523,18580..18940)
/locus_tag="CG11778"
/note="CG11778 gene product"
/codon_start=1

gene

map="45A1-45A1"
/db_xref="FLYBASE:FBgn0011746"
join(3648..4173,8618..8997,9387..9635,10420..10492,10917..12101)
/gene="ana"
/locus_tag="CG8084"
/product="CG8084-Ra"
/db_xref="FLYBASE:FBgn0011746"
join(13873..4173,8618..8997,9387..9635,10420..10492,10917..11343)
/gene="ana"
/locus_tag="CG8084"
/note="ana gene product"
/codon_start=1
/product="CG8084-PA"
/db_xref="GI:7303954"
/db_xref="FLYBASE:FBgn0011746"
/translation="MASAMRGKCRSRIRLVLILSLITMGADSRAPFPDPSFIEGVOSEVNPFRITLIRFNLEKQILISIQNRSNPMDDSAQSNQOYLQAVATORDIFRVOKALSNENPGASKEKAGFPICNAETTPEDMSIGNVTLOFASVPSINDDRLSSALIRIKTNPGOTREHNPQASTOPISTENGNTA PNCADPVPQIPQIVYSI VHQKKQKRTCTNTAMSSSTGVEIDYKCALAYEQQRQOQPOQOLTA SVGLIMEVHDEENLIRPGIFAPPTCDQADIAVPSVYTEPKSKLAWTLPRK PLDIFPNSNMKNSYNTPKSRAPLESTTSPITNDQSDSEYSSQQRVAPLVR HRRHHNHQCPSESESAQLEAEIEAEIEMSSASNSQOQMEPISNHRHRTGHHPHH OLOHHHHHRTKHKRIIPAKHQE"
13293..16922
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="last curated on Mon Feb 11 10:24:48 PST 2002"
/map="45A1-45A1"
/db_xref="FLYBASE:FBgn0025709"
join(13293..13488,14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/product="CG8083-Ra"
/db_xref="FLYBASE:FBgn0025709"
join(14663..14716,14775..14970,15024..15225,15288..16032,16117..16444,16512..16692,16754..16922)
/gene="BEST:CK00459"
/locus_tag="CG8083"
/note="BEST:CK00459 gene product"
/codon_start=1
/product="CG8083-PA"
/protein_id="AAP58997.1"
/db_xref="GI:730395

OY	1660	GGTGTGTGTTTAAAGAGCGCGCGCGTGTGTGTGTGCACATGGGAGATTAATCTAATCC	1719
Db	143998	GTGTGTGTGTTTAAAGAGCGCGCGCGTGTGTGTGTGCACATGGGAGATTAATCTAATCC	144057
OY	1720	GAAGTATATCTAAGTGTTTATAGTTCCTGTGTGGCGCAATATATCGAGCTAATTTAT	1779
Db	144058	GAAGTATATCTAAGTGTTTATAGTTCCTGTGTGGCGCAATATATCGAGCTAATTTAT	144117
OY	1780	GGCTCTGTATGATAGTTTGTCAATTGTGTGCAACCTCTGTGTGGCAGGCCCAAGAAA	1839
Db	144118	GGCTCTGTATGATAGTTTGTCAATTGTGTGCAACCTCTGTGTGTCCAGGCCCAAGAAA	144177
OY	1840	ATAAAGTTTTTTCATTGTGTGA	1863
Db	144178	ATAAAGTTTTTTCATTGTGTGA	144201

RESULT 4					
LOCUS	AC017981/c				
DEFINITION	AC017981	51430 bp	DNA	linear	HTG 09-DEC-1999
ACCESSION	AC017981				
VERSION	AC017981.1	GI:6553209			
KEYWORDS	HTG; HTGS_PHRASE2.				
SOURCE	Drosophila melanogaster				
ORGANISM	Drosophila melanogaster (fruit fly)				

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 51430)
Adams, M. and Venter, J. C.
Direct Submission
Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive

COMMENT This sequence was identified as CDM-10212888 by the submitter. For more information on this record e-mail to fly@celera.com.

*** NOTE:** This is a 'working draft' sequence.

*** This sequence will be replaced**

*** by the finished sequence as soon as it is available and**

*** the accession number will be preserved.**

Location/Qualifiers

```

BASE COUNT      13860 a 11923 c 11598 g 14049 t
ORIGIN
      /organism="Drosophila melanogaster"
      /mol_type="genomic DNA"
      /db_xref="taxon:7227"

```

Query Match	47.6%	Score 886;	DB 2;	length 51430;
Best Local Similarity	84.8%;	Pred.	No. 1.7e-185;	
Matches 1128; Conservative	0;	Mismatches	0;	Indels 202; Gaps 4;

QY	1	CGAGTCGAGCAAGGCAAGCAAGAAATTTTCAACCAACATTTTGGTTGGATGTTCAAGT	60
Db	1330	CGAGTCGAGCAAGGCAAGCAAGAAATTTTCAACCAACATTTTGGTTGGATGTTCAAGT	1271
QY	61	CACATTTTCGGTTGGTTTTAGTCCCTCTGCGCTCTCCCTGGTTCCTTACTTATGACATTT	120
Db	1270	CACATTTTCGGTTGGTTTTAGTCCCTCTGCGCTCTCCCTGGTTCCTTACTTATGACATTT	1211
QY	121	TTTCGTTCCCTTGTGTTTCATTTAAACCCCAACCAAGTAACGAAATCGAGGGCGCATGTG	180
Db	1210	TTTCGTTCCCTTGTGTTTCATTTAAACCCCAACCAAGTAACGAAATCGAGGGCGCATGTG	1151
QY	181	TGGAATTCGTCGATTCGCTGTGTCGTCATTTGGGCGCTGCCTGTGGGAGCAAAACCCGCTC	240
Db	1150	TGGAATTCGTCGATTCGCTGTGTCGTCATTTGGGCGCTGCCTGTGGGAGCAAAACCCGCTC	1091
QY	241	TTTTTGTGGGAGCCATTTTGTGGGAGCCATTTTGTGGGAGCCATTTTGTGGGAGCCATTTT	256
Db	1090	TTTTTGTGGGAGCCATTTTGTGGGAGCCATTTTGTGGGAGCCATTTTGTGGGAGCCATTTT	1031

QY	257	-----ACAGTGTGGCGAAACCTCCCTGGAACAGGTGTGCCAAGG	297
Db	1030	GTGCGCTTCTCTCTTTTTCACAGTGTGGCGAAACCTCCCTGGAACAGGTGTGCCAAGG	971
QY	298	GAGTTTCCGAGCAAGTTGAGTGCATGTCGTGAAGATCAATGTGTGTGAGGAA	357
Db	970	GAGTTTCCGAGCAAGTTGAGTGCATGTCGTGAAGATCAATGTGTGTGAGGAA	911
QY	358	AATGGC-----	363
Db	910	AATGGCGTTAAGTCTTGAAGAACCACTTAAGATGCTTAAGGCTTGTGACATGTGTGTGT	851
QY	354	-----CTGAGCAGCAAGGACTTCTGTGTCTTCAACTCCACAGCGCAGTCTTGC	411
Db	850	TCGCCAACTGACGTGAGCAGCAAGGACTTCTGTGTCTTCAACTCCACAGCGCAGTCTTGC	791
QY	412	TACGCCCAAGCTGACAGGAGTGAAGCCCAAGACTTACACAGGCTGGAAGAACCCCTGC	471
Db	790	TACGCCCAAGCTGACAGGAGTGAAGCCCAAGACTTACACAGGCTGGAAGAACCCCTGC	731
QY	472	GAGGCATCGGCTCAGTGTGGCGGCCAAGGCGCGAGCACTTCATGATGTGCGAGTGAAG	531
Db	730	GAGGCATCGGCTCAGTGTGGCGGCCAAGGCGCGAGCACTTCATGATGTGCGAGTGAAG	671
QY	532	CTGACCAAGCGCGGCCAAGTGTGTGCGTAGTACGCGCTGTGTTGTGACCTTGAGGACGCG	591
Db	670	CTGACCAAGCGCGGCCAAGTGTGTGCGTAGTACGCGCTGTGTTGTGACCTTGAGGACGCG	611
QY	592	GCCCAAGACCGCGGAGGCCAAGTCTGGAATGACAT-----	624
Db	610	GCCCAAGACCGCGGAGGCCAAGTCTGGAATGACATGAGTGGGTCAACAATACGTACTACAA	551
QY	625	-----GACGCGCGCAATGCT	639
Db	550	GAGCGTCAATTAAGACACTAATTTTAAACAATGATTTTAAACAATTTAAGACGCGCGCAATAGT	491
QY	640	GCTATCAGCAAGCAAGTTGAGTGTGAAGGTGCTTATCTGTACTGTGGCGCGCCCTTCCACG	699
Db	490	GCTATCAGCAAGCAAGTTGAGTGTGAAGGTGCTTATCTGTACTGTGGCGCGCCCTTCCACG	431
QY	700	GCTCCTGTGTGTGACGCGTGTACCGCGCGGTGACACCGCGCGCAACCGGAGGACATCATG	759
Db	430	GCTCCTGTGTGTGACGCGTGTACCGCGCGGTGACACCGCGCGCAACCGGAGGACATCATG	371
QY	760	TGGAAGTGACCAATCAAGTTTCAAGATCTTCTACACTGTGCTTCAACCGCAACCC	819
Db	370	TGGAAGTGACCAATCAAGTTTCAAGATCTTCTACACTGTGCTTCAACCGCAACCC	311
QY	820	ATCACCGTCAACGACCTCAAGCTCAACCACTTGAAGCTTACCAAGCTTACCTGTTATG	879
Db	310	ATCACCGTCAACGACCTCAAGCTCAACCACTTGAAGCTTACCAAGCTTACCTGTTATG	251
QY	880	GACACATGTGTGCGCGACCAAGCACTCACTTGTGACGTTGTTTAAATGAGCGATAC-TT	938
Db	250	GACACATGTGTGCGCGACCAAGCACTCACTTGTGACGTTGTTTAAATGAGCGATACCTT	191
QY	939	TAGTGTGAGCAATTTGTTGTTACAGACACAACTTCCGCTTCCAGCGCGGTGATGCCCC	998
Db	190	TAGTGTGAGCAATTTGTTGTTACAGACACAACTTCCGCTTCCAGCGCGGTGATGCCCC	131
QY	999	AACCAAGTTCCTTACTGTGTGCGCAACTCAACCTTGAGTTCGCGCGCGGTCAACAT	1055
Db	130	AACCAAGTTCCTTACTGTGTGCGCAACTCAACCTTGAGTTCGCGCGCGGTCAACAT	71
QY	1059	GTAACAACCCCTGAGCTTCAAGTCCCTGACAGCTGCAAGGCTCCCTTCATGTGTAAGTACAA	1118
Db	70	GTAACAACCCCTGAGCTTCAAGTCCCTGACAGCTGCAAGGCTCCCTTCATGTGTAAGTACAA	11
QY	1119	GGAGGACTTT 1128	
Db	10	GGAGGACTTT 1	

```

OY      1059  GTACAACACCCCTGAGCTTCAGAGTCCCTGACACTGACAGGCTCCCTTCAGATGACATACAA 1118
Db      70     GTACAACACCCCTGAGCTTCAGAGTCCCTGACACTGACAGGCTCCCTTCAGATGACATACAA 11
OY      1119  GGAGGACTTT 1128
Db      10     GGAGGACTTT 1

```

RESULT 5
AC018048/c
LOCUS AC018048 3678 bp DNA linear HTG 09-DEC-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
ACCESSION AC018048
VERSION AC018048.1 GI:6553142
KEYWORDS HTG; HTGS PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDM:10213188 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. .3678
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
BASE COUNT 999 a 848 c 788 g 1043 t
ORIGIN
Query Match 31.9%; Score 595; DB 2; Length 3678;
Best Local Similarity 100.0%; Pred. No. 4.3e-121; Indels 0; Gaps 0;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1269 CGATCCCAAGGCGAAGCTATACCATCAATGCGCGCGAGTAAACCGCTTACATCAT 1328
DB 3678 CGATCCCAAGGCGAAGCTATACCATCAATGCGCGCGAGTAAACCGCTTACATCAT 3619
QY 1329 CGGTTACACCTTATTTGATGTTGCAATGGCGACTCAATATTCATTTCTTATTCCTTCA 1388
DB 3618 CGGTTACACCTTATTTGATGTTGCAATGGCGACTCAATATTCATTTCTTATTCCTTCA 3559
QY 1389 GTCATCCACATGATATGTTACGAATCAATCTTATATACGATACGATTCACATA 1448
DB 3558 GTCATCCACATGATATGTTACGAATCAATCTTATATACGATACGATTCACATA 3439
QY 1449 ACACATTAACCGGAAATTAAGTCTCACAATTTGATGCTATCCGTAAATTTGCGCGCAA 1508
DB 3498 ACACATTAACCGGAAATTAAGTCTCACAATTTGATGCTATCCGTAAATTTGCGCGCAA 3439
QY 1509 GGATCAGTTTACGTTTCAGATCACAAGGCTGCTATCTTGTGCGCGCAAAATTTGCTTGA 1568
DB 3438 GGATCAGTTTACGTTTCAGATCACAAGGCTGCTATCTTGTGCGCGCAAAATTTGCTTGA 3379
QY 1569 GGTTCGTTGTCACGACGACGACATCCAGTATTTCAATTAGATAGCCCTTAATGTCGT 1628
DB 3378 GGTTCGTTGTCACGACGACGACATCCAGTATTTCAATTAGATAGCCCTTAATGTCGT 3319
QY 1629 TTATCGAAGAACTTTAAAGACGACATGCTGTGTGTTTAAAGACGCGCGCGTGTG 1688
DB 3318 TTATCGAAGAACTTTAAAGACGACATGCTGTGTGTTTAAAGACGCGCGCGTGTG 3259
QY 1689 TGTGTGTCACATGGCGAGTAATCTAATCCGAAGTATATCTAAAGTCTTATAGTTTCT 1748
DB 3258 TGTGTGTCACATGGCGAGTAATCTAATCCGAAGTATATCTAAAGTCTTATAGTTTCT 3199
QY 1749 TGTTCGCGCAATATGCGAATTAATTAAGCTCTGTATGATAGTATGTCATTTGTC 1808
DB 3198 TGTTCGCGCAATATGCGAATTAATTAAGCTCTGTATGATAGTATGTCATTTGTC 3139
QY 1809 AACCTCTGTTGTCGAGCGCCCAAGAAATTAAGTTTTCATTTGTTGA 1863

DB 3138 AACCTCTGTTGTCGAGCGCCCAAGAAATTAAGTTTTCATTTGTTGA 3084
RESULT 6
AX093993
LOCUS AX093993 445 bp DNA linear PAT 30-MAR-2001
DEFINITION Sequence 117 from Patent WO0118547.
ACCESSION AX093993
VERSION AX093993.1 GI:13510211
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS Davies, R.W., Kaiser, K.U. and Yang, M.Y.
TITLE Essential genes and assays relating thereto
JOURNAL Patent: WO 0118547-A 117 15-MAR-2001;
THE UNIVERSITY COURT OF THE UNIVERSITY OF GLASGOW (GB)
FEATURES
source
1. .445
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
misc_feature
1. .212
/note="Area matching Drosophila EST A1106957. This EST
forms a 1756bp contig with ESTs AA391125, AA567307 and
AA735971."
213..275
/note="Probable intron in gene represented by EST
A1106957"
276..382
/note="Area matching Drosophila EST A1106957"
BASE COUNT 81 a 105 c 121 g 138 t
ORIGIN
Query Match 13.1%; Score 244.6; DB 6; Length 445;
Best Local Similarity 82.7%; Pred. No. 1.7e-43;
Matches 320; Conservative 0; Mismatches 4; Indels 63; Gaps 1;
QY 45 GTTGTGATGTTCACTTTCGCTTGGTGTGTTTATGTCCTTCTGCTTCTGTTGTT 104
DB 1 GTTGTGATGTTCACTTTCGCTTGGTGTGTTTATGTCCTTCTGCTTCTGTTGTT 60
QY 105 CCTACTTATGCAATTTTGTCTCTGTTGTTTCAATTAACCCCAAGTAAGCGA 164
DB 61 CCTACTTATGCAATTTTGTCTCTGTTGTTTCAATTAACCCCAAGTAAGCGA 120
QY 165 ATCCAGCGCATGTTGGAATCGCTGATTTGCTGTTGCTGATTTGGGCTGCCGTGGC 224
DB 121 ATCCAGCGCATGTTGGAATCGCTGATTTGCTGTTGCTGATTTGGGCTGCCGTGGC 180
QY 225 GGAAGCAACGCCGCTCTTTTGTGGGAGCGCA----- 256
DB 181 GGAAGCAACGCCGCTCTTTTGTGGGAGCGCAAGGTGGTGAGCGCTGACCAAGTTG 240
QY 257 -----ACAGTGTGGCGAAACCTCCCTCGAA 281
DB 241 AATGTGATGTAAAGAGTGCCTTCTCTTTTTCACAGTGTGGCGAAACCTCCCTCGAA 300
QY 282 GACGATGTCCTCAAGTGAATTTGCGGACGATTTGCTGATTTGCTGGAAGATGATGAT 341
DB 301 GACGATGTCCTCAAGTGAATTTGCGGACGATTTGCTGATTTGCTGGAAGATGATGAT 360
QY 342 CGTGGCTTTCGAGGAAATGCTGAG 368
DB 361 CGTGGCTTTCGAGGAAATGCTGAG 387
RESULT 7
BTU10039
LOCUS BTU10039 2110 bp mRNA linear MAM 22-OCT-1994

DEFINITION Bos taurus vacuolar H+ ATPase subunit Ac45 mRNA, complete cds.
ACCESSION U10039
VERSION U10039.1 GI:497983
KEYWORDS Bos taurus (cow)
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 2110)
REFERENCE
AUTHORS Supek, F., Supekova, L., Mandiyan, S., Pan, Y. C., Nelson, H. and Nelson, N.
TITLE A novel accessory subunit for vacuolar H(+)-ATPase from chromaffin granules
JOURNAL J. Biol. Chem. 269 (39), 24102-24106 (1994)
MEDLINE 95014142
PUBMED 7929063
REFERENCE 2 (bases 1 to 2110)
AUTHORS Nelson, N.
TITLE Direct Submision
JOURNAL Submitted (25-MAY-1994) Nathan Nelson, Roche Institute of Molecular Biology, 340 Kingland St., Nutley, NJ 07110, USA
FEATURES
source
1. .2110
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/issue_type="adrenal medulla chromaffin granules"
/dev_stage="adult"
28. .1434
/note="vacuolar H+ ATPase subunit"
/codon_start=1
/product="Ac45"
/protein_id="AA50752.1"
/db_xref="GI:497984"
CDS
1132 TTTCGGGACCTCTGGGACCTGTGGGCTTCGTCGACGCCCGGATATCTGTAGGAGCTGTT 1191
1255 TTCTCTATGCTAGCCACCTGTGACGGCTCTTCTCCCGGATATCTGTAGGAGCTGCTC 1314
1192 GGTGGTGGCCCTACTGCTGCTGATCATGTTGCTGTGGGCTGCTGCTGATGATGACATCAAC 1251
1315 ACCTCTTGTTCATGCTGCTGATCTTCACTTACGAGGCTGACATGATCCTCAGCTCAAG 1374
1252 ACGATGACCGCTTTCGACGATCCCAAGGCGCAAGACTATCAACATCA 1297
1375 ACCATGATGCTTGCATGATGACCAAGGCGCCCAACATCACTTGA 1420
Db
RESULT 8
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS AF429315 Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 125020)
REFERENCE
AUTHORS Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S., Ingber, S.L., Ashworth, R.G., Fleisher, A., Stevanin, G., Brice, A., Potter, N.T., Ross, C.A. and Margolis, R.L.
TITLE A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2
JOURNAL Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE 21583737
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes, S.E., Ingber, S.L., Ashworth, R.G., Ross, C.A. and Margolis, R.L.
TITLE Direct Submision
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
source
1. .125020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/note="isolated from a patient with Huntington's disease-Like 2 (HDL2)"
/complement(35581..35746)
/rpt_type=tandem
/rpt_unit=ctg
/gene="JPH3"
/complement(436507..36887)
/note="synonym: JP3"
/complement(436507..36887)
/gene="JPH3"
/product="junctophilin 3"
/complement(436507..36887)
/note="component of the junctional complex between plasma membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AA40941.1"
/db_xref="GI:17646245"
CDS
1686 ATCGCARRMMKSHGRBRMYTSSKMSMMSVSYVMHSMHBSBQMBKCMTWS 16927
419 AGCTGACGAGATGAGGCCCAAGACTTACTACACAGCGTGAAGAACCTTGAGGAC 478
16928 CWSMNYKSSWMSGMCWGRSRKSKWMSRMSRMSRMSRMSRMSRMSRMSRMSRMSRMSR 16987
479 TCGGCTCAGTGGC--GCCAAGCGGAGACCACTCCATGATGATGATGATGATGATGATGAT 535
16988 YASMCNMCWSCCNRSCCNRSCCNRSCCNRSCCNRSCCNRSCCNRSCCNRSCCNRSCCNRSC 17047
536 CCACGCGGCGCAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 595
17048 KCMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMRSMR 17107
596 AGAGCCGAGGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
17108 MKMYSKMYCMYCMWGRSGYCSKMTSRMSRMSRMSRMSRMSRMSRMSRMSRMSRMSRMSR 17167
656 TCGAGTGAAGGTGCTTATCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 715

Db	17168	ACSTRGSSWSSKYCMRGSTYSTSGCCCTTTTTCCTCCCAANTGGGAAGCTTTTNCNKTY	17227
Qy	716	GTGCTACCCCGGTGACACCCGCGCAACACCGGTGATCATGTGGAAGTGCACCAATC	775
Db	17228	SYRKRRGACMKCKNNYNNNSWMSSCDAGMSCTKYRSMWT-----SMASYCMGSMYYC	17282
Qy	776	AGTTTCAGATCTTCTACACCTGCCCTGTCTTACACAGGCAACCCATCACCCTTACCGAC	835
Db	17283	SMBSMASRGMSWSWSYMMCKMRKSMRMS:CYMCKCMCMKCYCMNMSMRGMSYYMVASWS	17342
Qy	836	TCAAGCTCACCAACTCTAGCTCTACCAAGGTATCCGTTGTATGACACATCTGTGCGCG	895
Db	17343	SSRCCYCTTRCTCWSCKSSKCTKSYMMMRMS:KRMKMKMRKMRGWSGMSASBSCTKYKS-	17401
Qy	896	ACAAGCCAATCATCTTTGACGTTGTATTAATGAGCGGATCATTTAGCTTGAGCAATTTG	955
Db	17402	---MRGMSMSKCYRACGCMWMMKGYMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR	17458
Qy	956	TCTACGACACACACACTTCGCTCCGCTCCAGCGCGGTGAATGCCCAACAGCTTCTTACT	1015
Db	17459	CMKCYSRMSR:CMWMMKSYCASCRSCAMMSGTMYMKASYMCMSSYCTCWTCTSYMSYRCT	17518
Qy	1016	CGTCCGGCAACTCAACCTCTGAGTCCGCGCGCGCTCAACAATGTATCAACACCTTAGCT	1075
Db	17519	CMCGWMSYTKSKWSSSMSYKKGKRSYSWCTSRGANSWRRCTYRGSASMRAGSM	17578
Qy	1076	TCAAGTCCCTGCACACTCAGGCTCCCTTCATGAGTACGT--ACAAGAGAGACTTTCCTT	1133
Db	17579	SRRRAKGRSMGGRSKMMMTGTGMMMSKYYCTGRBMMMTTTCMMRBSYMYRSMASAG	17638
Qy	1134	CGGCGACTCTCGGACCTGTGTGGCTTCGTGAGCGCCCGGTATCCTGATGGACTGTTGT	1193
Db	17639	MRKSSWSGMRWMSGASSNRCKSASRSSWCSGRMKMGMS:CMSSMWGSSRSASCKG	17698
Qy	1194	GGTGGCCTTACTGCTGTCATCATGTTCTGCGGTCTGCGATGATGACATCAACAC	1253
Db	17699	SRGMRBRKRSKRYRKRGKGRKMTYSKSGSKGKCMWKSGSWTSCTYYTSSCMWMM	17754
Qy	1254	GATGAGCGCTTCACGATCCCAAGGGCAAGACTATACCATCATGATCCCGCGCGAGTA	1313
Db	17759	SKSKMCCMMMKRCKACCTYSMSCTSMYRCMGMSYSCTGTTCKCKTGKSKSRWY	17818
Qy	1314	AACCGCTCATCATCTCGGTTGAC	1336
Db	17819	WMSWTSKMYMWRRAACMYCYMS	17841
RESULT 9			
AK026519		1694 bp	linear
LOCUS			PRI 29-SEP-2000
DEFINITION			Homo sapiens CDNA: FLJ22866 f1s, clone KAT02286, highly similar to
ACCESSION			AK026519
VERSION			AK026519.1
KEYWORDS			oligo capping; f1s (full insert sequence).
SOURCE			Homo sapiens (human)
ORGANISM			Homo sapiens
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE			1 (bases)
JOURNAL			Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
AUTHORS			Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
TITLE			Nakamura,Y., Isogai,T. and Sugano,S.
JOURNAL			NEO human cDNA sequencing project
AUTHORS			Unpublished
TITLE			2 (bases 1 to 1694)
JOURNAL			Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
AUTHORS			Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE			Direct Submission
JOURNAL			Submitted (29-AUG-2000) Sunio Sugano, Institute of Medical Science,
AUTHORS			University of Tokyo, Laboratory of Genome Structure Analysis, Human
TITLE			Genome Center, Shitokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,

COMMENT
Japan (E-mail: cdna@ems.u-tokyo.ac.jp, Tel: 81-3-5443-5286, Fax: 81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry (cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' & 3' end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

```

FEATURES
source      Location/Qualifiers
1..1694
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="KATO2286"
            /cell_line="KATO III"
            /cell_type="signet-ring cell carcinoma"
            /clone_1ib="KAT"
            /note="cloning vector pMR18SFL3"
misc_feature 1..1694
            /note="highly similar to HDXQC Human mRNA for ORF"
BASE COUNT  346 a      507 c      429 g      412 t
ORIGIN

```

	Query March	3.2%;	Score 60.4;	DB 9;	Length 1694;
	Best Local Similarity	60.2%;	Pred. No. 0.013;		
	Matches 100;	Conservative	0;	Mismatches 66;	Indels 0;
				Gaps 0;	
QY	1132 TTCGGCGACTCCTCGGACCTGTGTGGGCTTCCTGACGCCCGGTATCTGATGGGACTGTT				1191
DB	853 TTCTCCTACGCGACGACTGTGCCAGCTTTCTTCCCGGACATGTGATGGGCTGCTC				912
QY	1192 GTGGTCGCCCTACTGCTGCTATCATGTTCGTGGCGCTCTGCTGATGATGGACATCAAC				1251
DB	913 ACCTCCCTGTTCATGCTCTTTCATCTTCAACCATGAGCCGACATGATCCTCAGCCTCAAG				972
QY	1252 ACGATGACGGGCTTCGACGATCCCAAGGGGCAAGACTATCACATCA				1297
DB	973 ACCATGATCGCTTGTGATGACCAAGGAGCCCACTATTTTCTTTGA				1018

RESULT 10	AX086813	LOCUS	DEFINITION	ACCESSION	VERSION
	AX086813	2043 bp	Sequence 765 from Patent WO0112659.	AX086813	AX086813.1 GI:13276038

```

KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
               Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS      Wiemann, S.
TITLE         Human dna sequences
JOURNAL       Patent: WO 0112659-A 765 22-FEB-2001;
               German Human Genome Project (DE)

FEATURES
source        Location/Qualifiers
               1..2043
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"

BASE COUNT    375 a      617 c      564 g      487 t

ORIGIN

```

	Query Match	Similarity	3.2%	Score	60.4	DB	6	Length	2043
	Best Local	Similarity	60.2%	Pred.	No.	0.013			
	Matches	100	Conservative	0	Mismatches	66	Indels	0	Gaps
Qy	1132	TTGCGGCACTCTCTGGGACTGTGTGGGCTTGTGACGCCCGGTATTCGTGATGGGACTGTT	1131						
Db	1229	TTCTCTCTACGCGAGGACTGTGGCAGGCTTCTTCTCCCGGCAATCTGGATGGGAGTGTCTC	1288						

OY		1192	GTGGTCGCCCTACACTGATCATATGTTCCGCAGCGTGTGGATCATGACATGAC	1251
Dd		1289	ACCTCCCTGTTTCAACCTCTTCACTTCACTTAAGCCTTGACATGATCTCAAGCTC	1348
OY		1252	ACGATGACCGCTTGACATGCCAAGGGCAAGACTATCACCATCA	1297
Dd		1349	ACCATGATCGCTTTGATGACACAAGGGCCCCCACTATTCTTCTTGA	1394
RESULT 11				
LOCUS		HSMB01819		
DEFINITION		HSM801819	2043 bp mRNA linear PRI 20-MAR-2002	
ACCESSION		AL136851		
VERSION		ALI36851.1	GI:12053206	
KEYWORDS		complete cds.		
SOURCE				
ORGANISM		Homo sapiens (human)		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE		1 (bases 1 to 2043)		
JOURNAL		Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S. Direct Submission Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de; Sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFPz434B0535) is available at the RZPD in Berlin. Please contact the RZPD: Resourcentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/ Location/Qualifiers		
FEATURES				
SOURCE				
gene				
CDS				
polyA_signal				
polyA_site				
BASE COUNT				

Query Match	3.2%	Score 60.4	DB 9	Length 2043
Best Local Similarity	60.2%	Pred. No. 0.013		
Matches 100	Conservative 0	Mismatches 66	Indels 0	Gaps 0
Qy	1132	TTTCGGCGGCTCTCTGGGAGCTGTGTGGGCTTCGTGACGCCCGGATCTGATGGAGCTGTT	1192	
Db	1229	TTTCCTTACGCGCAGAGCTGTGCGACGCTTCTTCTCCCGGATCTGATGGGCTGCTC	1288	
Qy	1192	GTGTTCGCCCTACTGCTGTGTCATGATGTTTCGTGGCGGCTGTCTGATGATGAGCATCAAC	1251	
Db	1289	ACCTCCCTGTTTCATGCTCTTTCATCTTTCATCTTTCATGAGCGCTGACATGATCCTCAGCCTCAAG	1348	
Qy	1252	ACGATGACAGCGCTTCGACGATATCCAGGGGCAAGACTATCACCATCA	1297	
Db	1349	ACCATGATGCTCTTGTGATGACCAAGGCGCCCACTATTCTTTGA	1394	
RESULT 12				
LOCUS	BD127635	2063 bp	DNA	linear
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.			
ACCESSION	BD127635			
VERSION	BD127635.1	GI:23222580		
KEYWORDS	JP 2002017375-A/3066.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 2063) Oda,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.			
TITLE	Primer for synthesizing full-length cDNA and use thereof			
JOURNAL	Patent: JP 2002017375-A 3066 22-JAN-2002;			
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002017375-A/3066 PD 22-JAN-2002 PF 07-JUL-2000 JP 2000253172 PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO ISHII, PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI SHINICHI KOJIMA, PI TETSUO OTSUKI,HISASHI KOGA PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10, PC C12P1/02,C12P1/68/C12P1/08,G06F1/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT CDS (26)..(1435). 1..2063 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" BASE COUNT 370 a 622 c 578 g 493 t ORIGIN			

	Db	1379	ACCATGATCGCTTGTATGACCACAGGGCCCACTATTTCCTTTGA	1424
RESULT 13				
LOCUS	AK075284			
DEFINITION	Homo sapiens cDNA FLJ90803 f18, clone Y79AA1000270, moderately similar to Bos taurus vacuolar H+ ATPase subunit Ac45 mRNA.			
ACCESSION	AK075284			
VERSION	AK075284.1			
KEYWORDS	oligo capping; f18 (full insert sequence).			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Bukariyota, Mataroa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Isegai,T., Oca,T., Nishikawa,T., Hayashi,K., Otsuki,T., Sugiama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai-Hio,Y., Satto,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Kojima,S., Nagahara,K., Masuno,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Niimuya,K.			
TITLE	NEDO human cDNA sequencing project			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 2063)			
AUTHORS	Isegai,T. and Otsuki,T.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-MAR-2002) Takao Isegai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 229-0812, Japan (E-mail:genom@helix.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)			
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.).			
FEATURES	Location/Qualifiers			
source	1..2063			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="Y79AA1000270"			
	/cell_line="Y79"			
	/cell_type="retinoblastoma"			
	/clone_id="Y79NA1"			
	/note="cloning vector: pME18SF13"			
CDS	26..1438			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="BAC11520.1"			
	/db_xref="gi:22761272"			
	/translation="MMAAMATARRMRGRCAQALMKRPMLEPVLISLAAAAAABEQDVPVLIVSSDRDLMAPADTHEGHVTSIDLPSTLYLDPALEIPRVLLFLDKISTEDPTAYGVFGNKSDSPASNLNMLADLPASSLVYPADWVAVSTLTLYLOEKLGASPLAVDTLATGEELKINSILPALILRLPYTASGLMAPREVLTGNDENVIGQSITKSDEVPYDAIAKTOMEDLPRLVARADVAVAGGIQLOKOVSPVTHPPVSINDAPRIIPMAONFESVAYKOWMEDLPRLTVRGVGLNLTGSFKMNSPARLSIYTRLEGCTVTRELIANLVYV SARHFPTMERLEVHSNGSVAFNMASQVTFSGSYSHCHCYVSLSKSGSLVARYOPSEPSARKHPFTMERLEVHSNGSVAFNMASQVTFSGSYSHCHCYVSLSKSGSLVARYOPSEPKTMMDIFDIQAIFNVMGQFVASDCASFSPGIMWGILLTSLFMFLFTYGHLMTSLDKTMDFDDHKHGKFTSLTIQIV"			
BASE COUNT	370 a 622 c 578 g 493 t			
ORIGIN				
Query Match	3.2%; Score 60.4; DB 9; Length 2063;			
Best Local Similarity	60.2%; Pred. No. 0.013;			
Matches 100; Conservative	0; Mismatches 66; Indels 0; Gaps 0;			
DB	1132 TTCGGCAGACTCCTGGGACATCTGTGGGCTTCGAGAGCCCGGATCCGATGGAGCTGTT	1191		
	1259 TTCTCTTAGCCAGGACATGTCAGACTTCTTCTCCCCCGGCATTGATGAGGCTGCTC	1318		

QY	1192	GTGTGGCCCTACTGCGTGCATCATGTTGCGGGGCTGCGAGATGAGCATCAAC	125
Db	1319	ACCTCCCTGTTCAAGCTCTTCACTTTCACCTATAGCGCTGCACATGATCTTCAGCTCAAG	1378
QY	1252	ACGATGAGCCGCTTCGACGATCCCAAGGGCAAGCTATCACTCA	1297
Db	1379	ACCATGATGCTTGTATGATGACCAAGGGCCCACTATTTCTTTGA	1424
RESULT 14			
AK091083			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
Query Match			
Best Local Similarity			
Matches			
QY	1132	TTTGGCGAGCTCTGGGACTGTGTGGGGCTTGGACGCCGGTATCTGATGGAGCTGTT	1191
Db	1276	TTTCTCTTACGCGAAGGACTGTGTGCCAGCTTTCTTCCCGGCACTGTGATGGGGTGTCTC	1335
QY	1192	GTGTGTGCGCCCTACTGCTGTGTCATCATGTTTGTGTGGCGGTCTGTGATGATGAGCATCAAC	1251
Db	1336	ACCTCCCTGTTCAAGCTCTTTCATCTTTCACCTATGAGCGCTGCACATGATCTTCAAGCTTCAAG	1395
QY	1252	ACGATGAGCCGCTTCGACGATCCCAAGGGCAAGCACTATCACTCA	1297

Db 1396 ACCATGATCGCTTGTATGACCAAGGGCCCACTATTCTTTGA 1441

RESULT 15
BC000724 2087 bp mRNA linear PRI 12-JUL-2001

LOCUS Homo sapiens, Similar to ATPase, H+ transporting, lysosomal
(vacuolar proton pump), subunit 1, clone MGC:1570 IMAGE:3506925,
mRNA, complete cds.

ACCESSION BC000724 GI:12653864

VERSION BC000724.1

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LIML)

<http://www.systemsbio.org>

Contact: amadan@systemsbiology.org

Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAL Plate: 7 Row: 9 Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4557340.

Location/Qualifiers

1. .2087

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:1570 IMAGE:3506925"

/issue_type="Kidney, renal cell adenocarcinoma"

/clone_lib="NIH MGC_14"

/lab_host="DH10B-R"

/note="Vector: POTB7"

28. .1440

/codon_start=1

/product="Similar to ATPase, H+ transporting, lysosomal
(vacuolar proton pump), subunit 1"

/protein_id="AAH00724.1"

/db_xref="GI:12653865"

/translation="MMAAMATARVRMGPRCAQLMMPMLPVFLSLAAAAAAAEQ
VPLVWSSDRDLWAPADTHGHITSDLOSTYLDPALEGPNNVLLFLQDKLSIED
TAYGVFGNKODSAPGNLEALDLAPSSIVLPADWVAVSTLTLYOEKLGASPLVD
LATRELKLNASLPEALLIRLPTASSGMAPREVLTGNDEVITGOVLTLSKSEDPYT
AALTAVPRSVADVAVAGGILROKOPVSPVTHPPVSNDTAPRILFMANQNSV
AYKQWEDLTPLTFVGOELNLTGSPFNDSFARSLTYERLFGTTVPKFLANRLYPV
SAHWPTMERLEVHNSGVAVFNASOVTPGSIYFPCYVSSLSKSGSLIVARTQSP
WQMLQDPQIQAFNVMGEOFSYASDCASPPSPGIMMGLLTSLFMLPIFTYGLMILSL
KTMDFRDPDHKGPTISLTQIV"

BASE COUNT 397 a 620 c 578 g 492 t

ORIGIN

Query Match 3.2%; Score 60.4; DB 9; Length 2087;
Best Local Similarity 60.2%; Pred. No. 0.013;
Matches 100; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 1132 TTGGGAGCTCTGGAGCTGTGTGGCTTGTGACGCCCGGTATCTGATGGAGCTGTT 1191
Db 1261 TTCTCCTACGACGAGCACTGTGCAGCTTCTCTCCCGGACATGTGATGGGCTGCTC 1320
Qy 1192 GTGGTGGCCCTACTGTGTGATCATGTTGTGTGGGTCTGTGTGATGATGATCATCA 1251
Db 1321 ACCTCCCTGTTCATGCTCTTCACTTCACTTATGGCTGCACATGATCTTCAAGCTCAAG 1380
Qy 1252 ACCATGACCGCTTGCAGCATCCCAAGGCGAAGACTATGACCATCA 1297
Db 1381 ACCATGATCGCTTGTATGACCAAGGCGCCCACTATTCTTTGA 1426

Search completed: February 8, 2004, 04:40:33
Job time : 7109.72 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: February 7, 2004, 18:33:25 ; Search time 530.793 Seconds
(without alignments)
9474.605 Million cell updates/sec

Title: US-09-614-150A-5

Perfect score: 1863

Sequence: 1 cgacgtcgagcaagcaag.....agtttttccaatgtgtta 1863

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1863	100.0	1863	23	ABLO1843
2	1691	90.8	1758	23	ABLO3059
3	1632	87.6	4064	23	ABLO1842
4	1632	87.6	4064	23	ABLO3058
5	268	14.4	4089	23	ABLO4554
6	244.6	13.1	445	22	AAH28928
7	60.4	3.2	2037	24	AA562238
8	60.4	3.2	2043	23	ABX71394

9	60.4	3.2	2063	22	AAK94606
10	60.4	3.2	2156	20	AAK81298
11	60.4	3.2	2823	24	ABK83610
12	59	3.2	7637	23	ABLO3844
13	55.6	3.0	4416	23	ABLO3845
14	54	2.9	40116	25	AB226080
15	51.2	2.7	8244	23	AA566529
16	50.6	2.7	985	24	ABO44816
17	50.6	2.7	985	24	ABO44817
18	48.4	2.6	1590	24	AA518432
19	47.8	2.6	2400	25	ACC44557
20	47.8	2.6	2478	25	ACC44567
21	47.6	2.6	1543	23	AA592206
22	46.4	2.5	1973	22	ABA44805
23	46.4	2.5	1973	22	ABA55261
24	46.4	2.5	1973	22	ABA25005
25	46.4	2.5	1973	22	AAK03514
26	46.4	2.5	1973	22	AAK28970
27	46.4	2.5	1973	22	AAI13556
28	46.4	2.5	1973	22	AAI34918
29	46.4	2.5	1973	22	AAI03446
30	46.4	2.5	1973	23	AB228585
31	46.4	2.5	1973	24	AB503504
32	45.4	2.4	10732	21	AAAI0594
33	45.4	2.4	112460	24	ABK83567
34	45.4	2.4	349980	24	ABO81848
35	45	2.4	519	22	AA501997
36	44.8	2.4	349980	24	ABO81846
37	44.2	2.4	1065	25	ABT32132
38	44.2	2.4	32539	25	ABT32129
39	44	2.4	2037	25	ACC44573
40	43.2	2.3	1374	22	AA54487
41	43.2	2.3	1944	22	ABA44123
42	43.2	2.3	1944	22	AB545475
43	43.2	2.3	1944	22	AB243359
44	43.2	2.3	1944	22	AAI12874
45	43.2	2.3	1944	22	AAI34231

ALIGNMENTS

RESULT 1	
ID	ABLO1843
strand	standard; cDNA; 1863 BP.
AC	ABLO1843;
DT	26-MAR-2002 (first entry)
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 11.
XX	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ss.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
XX	
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656660/75.
XX	
DR	P-PSDB; ABB57740.
XX	

Human full-length
Human vacuolar ATP
Human cDNA differe
Drosophila melanog
Drosophila melanog
Mouse Dnaex encod
DNA encoding novel
Oligonucleotide fo
Contig 143a DNA en
Maize 6gpi pullula
Plasmid pMOV7005 6
DNA encoding novel
Human breast cell
Human foetal liver
Probe #3471 for ge
Human bone marrow
Human brain expres
Probe #3489 for ge
Probe #3604 used t
Probe #3437 used t
Human liver single
Human genome-deriv
Gene encoding a su
Human cDNA differe
Bifidobacterium 10
Banana phosphoenol
Bifidobacterium 10
Benzodiazepines bi
Benzodiazepines bi
Glucosylase nucle
Botulism toxin hea
Human breast cell
Human foetal liver
Probe #2825 for ge
Probe #2807 used t
Probe #2917 used t

Db 1801 ATTTGTCACCTCTCTGTTGTCGAGGCCCAAGAAATTAAGTTTTCATTTG 1860

QY 1861 TTA 1863

Db 1861 TTA 1863

RESULT 2

ABL03059

ID ABL03059 standard; cDNA; 1758 BP.

XX ABL03059;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 3659.

DE Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical; gene; ss.

XX Drosophila melanogaster.

OS

XX MO200171042-A2.

PN

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

XX

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PMD, Myers EW;

XX

DR WPI; 2001-656860/75.

DR P-PSDB; ABB58956.

XX

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX

PS Claim 1; SEQ ID NO 3659; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-AB16175) and the encoded proteins

CC (AB5737-AB572072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

XX

SO Sequence 1758 BP; 396 A; 492 C; 423 G; 447 T; 0 other;

Query Match 90.8%; Score 1691; DB 23; Length 1758;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1691; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 CGATGTTGGAATATGCTGATTTGCTGCTCATTTGGGGTGGCCGTGGAGACAA 232

Db 68 CGATGTTGGAATATGCTGATTTGCTGCTCATTTGGGGTGGCCGTGGAGACAA 127

QY 233 CGCCGCTTTTGTGGGGAGCCAAAGTGTGCGAAACCTCCCTGAAGACGGTGTCC 292

Db 128 CGCCGCTTTTGTGGGGAGCCAAAGTGTGCGAAACCTCCCTGAAGACGGTGTCC 187

QY 293 AAGTGAGTTTGGCGAGAGTTGGCTGATTTGCTGGAAGATCATGATGTCGTCCTGG 352

Db 188 AAGTGAGTTTGGCGAGAGTTGGCTGATTTGCTGGAAGATCATGATGTCGTCCTGG 247

QY 353 AGAAAAATGGCTTGAAGCAAGAACTTCTGTGCTCAACTCCAGGGCAGTCTGCT 412

Db 248 AGAAAAATGGCTTGAAGCAAGAACTTCTGTGCTCAACTCCAGGGCAGTCTGCT 307

QY 413 ACGCCAGTTCAGAGGAGTGAACCCCAAGACTTACTACACGAGCTGGAAGAACCTCTCG 472

Db 308 ACGCCAGTTCAGAGGAGTGAACCCCAAGACTTACTACACGAGCTGGAAGAACCTCTCG 367

QY 473 AGGCACTGCTCAGTGGCCGCGCAAGCGGAGCACACTTCATGCAATGCAATGGAAGAC 532

Db 368 AGGCACTGCTCAGTGGCCGCGCAAGCGGAGCACACTTCATGCAATGCAATGGAAGAC 427

QY 533 TGACCAAGCCGCGCAAGTGTGCTGAGTACGCGCTGTTTGTGAACCTTGAAGACGCG 592

Db 428 TGACCAAGCCGCGCAAGTGTGCTGAGTACGCGCTGTTTGTGAACCTTGAAGACGCG 487

QY 593 CCGAGAGCCGGAAGGCGAGTGTGAGTCAATGAGCGCGCATGAGTGTATGAGCAGC 652

Db 488 CCGAGAGCCGGAAGGCGAGTGTGAGTCAATGAGCGCGCATGAGTGTATGAGCAGC 547

QY 653 AGTTGAGTGAAGGCTTATCTGTACCTGCGCGCCCTCCAGCGCTCTGAGTGC 712

Db 548 AGTTGAGTGAAGGCTTATCTGTACCTGCGCGCCCTCCAGCGCTCTGAGTGC 607

QY 713 AGCGTGTACCGCGCGTGAACCGCGCGCAACCGGTGAGTCAATGTAAGTGCAGCA 772

Db 608 AGCGTGTACCGCGCGTGAACCGCGCGCAACCGGTGAGTCAATGTAAGTGCAGCA 667

QY 773 ATCAAGTTTCAATCTTCTTACACTGCTGCTGTACCAAGCGCAACCCCATCACCGTACCG 832

Db 668 ATCAAGTTTCAATCTTCTTACACTGCTGCTGTACCAAGCGCAACCCCATCACCGTACCG 727

QY 833 ACTCAAGTCAACCACTTACTCTTACCAACTTACCTGTTATGAGCAATCTGTG 892

Db 728 ACTCAAGTCAACCACTTACTCTTACCAACTTACCTGTTATGAGCAATCTGTG 787

QY 893 CCGAAGCGCAATCACTTGTGAGTGTATTAATGAGCGATTAATGAGTGTGAGCAAT 952

Db 788 CCGAAGCGCAATCACTTGTGAGTGTATTAATGAGCGATTAATGAGTGTGAGCAAT 847

QY 953 TGGTCTACGACAACAACAATTCGCTCCAGCGGCGGTGAATGCCCAACCAAGTCTCTT 1012

Db 848 TGGTCTACGACAACAACAATTCGCTCCAGCGGCGGTGAATGCCCAACCAAGTCTCTT 907

QY 1013 ACTCGTGGCAACTCACTTCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1072

Db 908 ACTCGTGGCAACTCACTTCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 967

QY 1073 GCTTCAAGTCCCTGAGGCTGAGGCTCCCTTGATGTGATGTAAGAGGAGCTTTCCT 1132

Db 968 GCTTCAAGTCCCTGAGGCTGAGGCTCCCTTGATGTGATGTAAGAGGAGCTTTCCT 1027

QY 1133 TCGGAGACTCTCGGAGACTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1192

Db 1028 TCGGAGACTCTCGGAGACTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1087

QY 1193 TGGTGCCCTTACTGCTGCTGATCATGTTCTGTGCGGTCTGTGAGTGTGAGTGTGAGTGTG 1252

Db 1088 TGGTGCCCTTACTGCTGCTGATCATGTTCTGTGCGGTCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1147

QY 1253 CGATGAGCCGCTTGAACGATTCGCAAGGAGCAAGTATCAACATGATGCGGCGGAGT 1312

Db 1148 CGATGAGCCGCTTGAACGATTCGCAAGGAGCAAGTATCAACATGATGCGGCGGAGT 1207

QY 1313 AAACGCGCTTACATGAGTGTGACACTTATTTGATTTGATGAGTGTGAGTGTGAGTGTGAGTGTG 1372

Db 1208 AAACGCGCTTACATGAGTGTGACACTTATTTGATTTGATGAGTGTGAGTGTGAGTGTGAGTGTG 1267

QY 1373 TTCTTTATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1432

Db 1268 TTCTTTATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1327

QY 121 TTGCTTCTCTTTGTTTTCATTAAACCCCGAAGTAAGCATCAGCGGATGTTG 180
DB 2944 TTGGTCTCTTTGTTTTCATTAAACCCCGAAGTAAGCATCAGCGGATGTTG 2885
QY 181 TGGAAATCGCTGATTTGGTGTGTGCTATTGGGGCTGCGGTGGAGCAAGCCCGTC 240
DB 2884 TGGAAATCGCTGATTTGGTGTGTGCTATTGGGGCTGCGGTGGAGCAAGCCCGTC 2825
QY 241 TTTTGTGGGAGGCA----- 256
DB 2824 TTTTGTGGGAGGCAAGCTGGGTGAGCGCCCTGCACAGTTGAATGTGATACGA 2765
QY 257 -----ACAGTGTGCGCAAAACCTCCCTGAAGACGATGCCCAAGT 297
DB 2764 GTGCTTCTCTCTTTTTCAGATGTGCGCAACCTCCCTGAAGACGATGCCCAAGT 2705
QY 298 GAGTTTTCGAGACAGTTGGCTGCTATTCTGGAGATCAATGCTGCTTGCAGAGAA 357
DB 2704 GAGTTTTCGAGACAGTTGGCTGCTATTCTGGAGATCAATGCTGCTTGCAGAGAA 2645
QY 358 AATGGC----- 363
DB 2644 AATGGCTTAAGCTTGAAGACCACTTAAGTACCTAGGCTTGTGACATGTGTTGT 2585
QY 364 -----CTGACGACGCAAGGACTTCTGTGTCTCAACTCCGAGCGCACTGTC 411
DB 2584 TCCCAACTGAGCTGAGCAAGAGGACTTCTGTGTCTCAACTCCGAGCGCACTGTC 2525
QY 412 TAGCCCAAGCTGACGAGGAGTGAAGCCCAAGACTTAACACAGCGTGGAGAACCTTCG 471
DB 2524 TAGCCCAAGCTGACGAGGAGTGAAGCCCAAGACTTAACACAGCGTGGAGAACCTTCG 2465
QY 472 GAGGCACTGGGCTCAGTGGCGCGCAAGGAGAGCAACTCCATCGATGTGACGAGGAAG 531
DB 2464 GAGGCACTGGGCTCAGTGGCGCGCAAGGAGAGCAACTCCATCGATGTGACGAGGAAG 2405
QY 532 CTGACCAAGCGCGCAAGTGTGCGGTAGTACGCGCTGTGTGTGACCTTGAAGAGCGC 591
DB 2404 CTGACCAAGCGCGCAAGTGTGCGGTAGTACGCGCTGTGTGTGACCTTGAAGAGCGC 2345
QY 592 GCCGAGAGCGCGGAGGCGCAAGCTGTGAGTCACT----- 624
DB 2344 GCCGAGAGCGCGGAGGCGCAAGCTGTGAGTCACTGTGGGTCAACATTAAGTACTACAA 2285
QY 625 -----GACGCGGCAATAGCT 639
DB 2284 GCGGTCAATTAAGACACACTAATTTAACAATGATTTAAACAATTTAAGCGCGCAATAGCT 2225
QY 640 GCTATCAGCAAGAGCTTGAAGTGAAGGTGCTTATCTGTAACCTGCGCGCCCTCCAGC 699
DB 2224 GCTATCAGCAAGAGCTTGAAGTGAAGGTGCTTATCTGTAACCTGCGCGCCCTCCAGC 2165
QY 700 GCTCTGTGTGACGCGTCTGTAACCGCGTGAACCGCGCGCAACCGGTGCGATCATG 759
DB 2164 GCTCTGTGTGACGCGTCTGTAACCGCGCGTGAACCGCGCGCAACCGGTGCGATCATG 2105
QY 760 TGGAAAGTGAAGCAATGCTTCAAGTCTTCACTGCACTGCTGCTTCAACGGAAGAGCC 819
DB 2104 TGGAAAGTGAAGCAATGCTTCAAGTCTTCACTGCACTGCTGCTTCAACGGAAGAGCC 2045
QY 820 ATGACGCTGACGCACTCAAGCTCAACCACTGAGCTTCAACAGCTATCGTGTGATG 879
DB 2044 ATGACGCTGACGCACTCAAGCTTCAACCACTGAGCTTCAACAGCTATCGTGTGATG 1995
QY 880 GACACATCTGTGCGCGCAAGCAAGCAATCACTTTGAGCTTGTATTAATGCGGATACTTT 939
DB 1984 GACACATCTGTGCGCGCAAGCAAGCAATCACTTTGAGCTTGTATTAATGCGGATACTTT 1925
QY 940 AGTCTGAGCAATTTGTGCTACGACAAACAACCTTCCGCTTCAAGCGCGTGAATGCCCA 999
DB 1924 AGTCTGAGCAATTTGTGCTACGACAAACAACCTTCCGCTTCAAGCGCGTGAATGCCCA 1865
QY 1000 ACCAGTTCCTTACTGTGTGCGCAACCTCACTGAGTCCGCGCGCTCAACCAATG 1059

DB 1864 ACCAGTTCCTTACTGTGTGCGCAACCTCACTGAGTCCGCGCGCTCAACCAATG 1805
QY 1060 TACAAACCTGAGCTTCAAGTCCCTGACGCTCAGAGCTCCCTCGATGATGATCAAG 1119
DB 1804 TACAAACCTGAGCTTCAAGTCCCTGACGCTCAGAGCTCCCTCGATGATGATCAAG 1745
QY 1120 GAGACCTTCCCTTCCGCGCACTCTCTGGAAGTGTGTGGGCTTCTGTAAGCGCGGATCTG 1179
DB 1744 GAGACCTTCCCTTCCGCGCACTCTCTGGAAGTGTGTGGGCTTCTGTAAGCGCGGATCTG 1685
QY 1180 ATGGACTGTTTGTGTGCTGCTTACCTGCTGCTGATATGTTGTTGCGGCTTCTGATG 1239
DB 1684 ATGGACTGTTTGTGTGCTGCTTACCTGCTGCTGATATGTTGTTGCGGCTTCTGATG 1625
QY 1240 ATGGAATCAACAGATGAGACGCTTGAAGATCCCAAGGAGCAACTATCACTATCAAT 1299
DB 1624 ATGGAATCAACAGATGAGACGCTTGAAGATCCCAAGGAGCAACTATCACTATCAAT 1565
QY 1300 GCCCGCGGAGTAAACCGCTTACATCATCGTTTCAACCTTATTTGATTTGCAATGCG 1359
DB 1564 GCCCGCGGAGTAAACCGCTTACATCATCGTTTCAACCTTATTTGATTTGCAATGCG 1505
QY 1360 CACTCAATATTCATTCCTTATTCCTTCACTGCTATCACTGATATATGTTAAGATTCAA 1419
DB 1504 CACTCAATATTCATTCCTTATTCCTTCACTGCTATCACTGATATATGTTAAGATTCAA 1445
QY 1420 TACTTAATACCGTACACGATTAACATTAACATTAACGAGAAATTAAGCTTCACAAT 1479
DB 1444 TACTTAATACCGTACACGATTAACATTAACATTAACGAGAAATTAAGCTTCACAAT 1385
QY 1480 TGTATGTAATCCGTAATTTTCCGCGCAAGATCAGTTTACGTTACGATCAAGCGTTG 1539
DB 1384 TGTATGTAATCCGTAATTTTCCGCGCAAGATCAGTTTACGTTACGATCAAGCGTTG 1325
QY 1540 CTGATTCCTGTGCGCGCAACAAATTCCTGAGGTTTCTTGCACCGACGACATCCAGTA 1599
DB 1324 CTGATTCCTGTGCGCGCAACAAATTCCTGAGGTTTCTTGCACCGACGACATCCAGTA 1265
QY 1600 TTTCAATTAAGATGAGCGCCCAATGTCGTTTATGGAAGAACTTAAAGACGACATGT 1659
DB 1264 TTTCAATTAAGATGAGCGCCCAATGTCGTTTATGGAAGAACTTAAAGACGACATGT 1205
QY 1660 GTGTGTGTGTTAAGAGCGCGCTGTGTGTGTGCAATGAGGAGTAAATCTAATCC 1719
DB 1204 GTGTGTGTGTTAAGAGCGCGCTGTGTGTGTGCAATGAGGAGTAAATCTAATCC 1145
QY 1720 GAAGTATATCTAAGTGTGTTATGTTTCTGTTTGGCGCAATATGCGAGACTAATTTAT 1779
DB 1144 GAAGTATATCTAAGTGTGTTATGTTTCTGTTTGGCGCAATATGCGAGACTAATTTAT 1085
QY 1780 GCTCTGTATGATGATGTTGCTATTTGGAACCTCTGTTGTTCCGAGGCGCAAGAA 1839
DB 1084 GCTCTGTATGATGATGTTGCTATTTGGAACCTCTGTTGTTCCGAGGCGCAAGAA 1025
QY 1840 ATAAAGTTTTCATTTGATTA 1863
DB 1024 ATAAAGTTTTCATTTGATTA 1001

RESULT 5
ABL04554/c
ID ABL04554 standard; cDNA; 4089 BP.
XX ABL04554;
XX AC
XX DT
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 8144.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX

OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
XX 27-SEP-2001.
PD
PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR P-PSDB; ABB60451.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 8144; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/publishe_pct_sequences.
XX
SQ Sequence 4089 BP; 1036 A; 992 C; 976 G; 1085 T; 0 other;
XX
Query Match 14.4%; Score 268; DB 23; Length 4089;
Best Local Similarity 84.4%; Pred. No. 6.8e-62;
Matches 341; Conservative 0; Mismatches 0; Indels 63; Gaps 1;
XX
QY 1 CGACGTCGAGCAAGCAAGCAAAATTTTACACACATTTTCTGTTAGTTGAGTT 60
DB 404 CGACGTCGAGCAAGCAAGCAAAATTTTACACACATTTTCTGTTAGTTGAGTT 345
QY 61 CACTTTTCGTTGGTTTAAAGTCCTTCTGTTCTTCTTCTTACTTATGCAATTT 120
DB 344 CACTTTTCGTTGGTTTAAAGTCCTTCTGTTCTTCTTCTTACTTATGCAATTT 285
QY 121 TTGGTCTCCTTTGTTTCAATTAACCCCAAGGTAAGCAATCCAGCGGAGTTG 180
DB 284 TTGGTCTCCTTTGTTTCAATTAACCCCAAGGTAAGCAATCCAGCGGAGTTG 225
QY 181 TGGAAATCGCTGATTCGTTGTCGTCATTTGGGCTGCGTGGAGCAAAAGCCCGTC 240
DB 224 TGGAAATCGCTGATTCGTTGTCGTCATTTGGGCTGCGTGGAGCAAAAGCCCGTC 165
QY 241 TTTTGTGGGAGCA-- 256
DB 164 TTTTGTGGGAGCAAGCAAGTGGTGAAGCGCTGACAGTTGAATGATGTAACGA 105
QY 257 -----ACAGTGTGGGCAAACTCTCCCTGAAGACGGTGTCCCAAGTG 297
DB 104 GTGCTTCTCTCTTTTTCACAGTGTGGGCAAACTCTCCCTGAAGACGGTGTCCCAAGTG 45
QY 298 GAGTTTCCGAGCAGTTGGCTGATTCGTTGAGTGAAGATCAATGAT 341
DB 44 GAGTTTCCGAGCAGTTGGCTGATTCGTTGAGTGAAGATCAATGAT 1

RESULT 6
AAH28928

ID AAH28928 standard; DNA; 445 BP.
XX
XX
AC AAH28928;
XX
XX 17-JUL-2001 (first entry)
DT
XX
XX Drosophila melanogaster essential gene fragment, SEQ ID NO: 117.
DE
XX
XX Drosophila melanogaster; fruit fly; essential gene; screening assay;
KM pesticide; crop protection; chromosome 2; ds.
XX
XX Drosophila melanogaster.
OS
PN WO200118547-A1.
XX
XX 15-MAR-2001.
PD
XX
XX 06-SEP-2000; 2000WO-GB03444.
PF
XX
XX 07-SEP-1999; 99GB-0021009.
PR
XX
XX (UNITV) UNITV GLASGOW.
PA
XX
XX Davies RW, Kaiser K, Yang MY;
PI
XX
XX WPI; 2001-281436/29.
DR
XX
XX Screening assays for used for identifying compounds having a
PT physiological effect on proteins identified as being essential -
PT
XX
XX Claim 1; Page 195; 695pp; English.
XX
XX The present sequence is part of an essential gene from Drosophila
CC melanogaster. Lack of expression of the protein encoded by this
CC gene leads to a lethal or semi-lethal phenotype. The invention
CC relates to 902 nucleic acid sequences from genes encoding proteins
CC which are thought to be essential, and to a screening assay for
CC identifying compounds which have a physiological effect on these
CC proteins. Suitable compounds are useful as pesticides and may be used
CC in conjunction with other pesticides and herbicides for crop
CC protection. The gene corresponding to the present sequence is located
CC on chromosome 2.
XX
SQ Sequence 445 BP; 81 A; 105 C; 121 G; 138 T; 0 other;
XX
Query Match 13.1%; Score 244.6; DB 22; Length 445;
Best Local Similarity 82.7%; Pred. No. 4.8e-56;
Matches 320; Conservative 0; Mismatches 4; Indels 63; Gaps 1;
XX
QY 45 GTTGTAGTTTCACTTCTGTTGGTTTAAAGTCCTTCTGTTCTTCTTCTTCTTCTT 104
DB 1 GTTGTAGTTTCACTTCTGTTGGTTTAAAGTCCTTCTGTTCTTCTTCTTCTTCTT 60
QY 105 CCTTACTTATGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 164
DB 61 CCTTACTTATGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 120
QY 165 ATTCAGCGCGATTTGTGAAATTCGTCATTCGTTGTCGTCATTTGGGCTGCGTGGC 224
DB 121 ATTCAGCGCGATTTGTGAAATTCGTCATTCGTTGTCGTCATTTGGGCTGCGTGGC 180
QY 225 GGAGCAAAAGCCCGCTTTTGTGGGAGCA-- 256
DB 181 GGAGCAAAAGCCCGCTTTTGTGGGAGCAAGTGGTGAAGCGCTGACCAAGTTG 240
QY 257 -----ACAGTGTGGGCAAACTCTCCCTGA 281
DB 241 AATGTAGTGTACAGATGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 300
QY 282 GACGTTTCCCAAGTGTGGTTTCCGAGCAGTTGGCTGATTCGTTGAGTGAAGATCAATGAT 341
DB 301 GACGTTTCCCAAGTGTGGTTTCCGAGCAGTTGGCTGATTCGTTGAGTGAAGATCAATGAT 360

QY 342 CGTGGCCTTCGAGAGAAATGCGCTGAG 368
|||
Db 361 CGTGGCCTTCGAGAGAAATGCGCTGAG 387

RESULT 7

AAS62238
ID AAS62238 standard; cDNA; 2037 BP.

XX AAS62238;

XX 14-FEB-2002 (first entry)

XX cDNA sequence #25 encoding novel human secreted protein.

XX Human secreted protein; hyperproliferative disorder; autoimmune disorder;
XX immune deficiency disorder; blood disorder; inflammatory disorder;
XX infectious disorder; gene therapy; antimicrobial; hepatotropic;
XX immunosuppressive; antirheumatic; ss.

XX Homo sapiens.

XX MO200177291-A2.

XX 18-OCT-2001.

XX 29-MAR-2001; 2001MO-US10485.

XX 06-APR-2000; 2000US-195604P.

XX (GEMV) GENETICS INST INC.

XX Wong GG, Clark HF, Fecthel K, Agostino MJ, Howes SH, Resnick RJ;
XX Galunota K, Graham JR;

XX WPI, 2002-010900/01.

XX New polynucleotides encoding secreted proteins useful for treating e.g.
XX asthma, HIV and Crohn's disease -

XX Claim 1; Page 84; 391pp; English.

XX The present invention relates to the isolation of novel cDNA sequences
XX which encode human secreted proteins. The cDNA sequences have been
XX derived from a variety of human tissues. The invention also provides
XX a method for producing proteins from these polynucleotide sequences.
XX The proteins are useful for identifying compounds that modulate their
XX activity and production, and the cell is also useful for identifying
XX compounds that modulate expression of the polynucleotide sequences
XX encoding the secreted proteins. The sequences of the invention are
XX useful for treating diseases such as hyperproliferative disorders
XX (e.g. cancer), immune deficiency disorders (e.g. severe combined
XX immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
XX sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory
XX disorders (e.g. arthritis) and infectious disorders (e.g. hepatitis).
XX The polynucleotide sequences of the invention are also useful in gene
XX therapy. AAS62234-AAS62838 represent the cDNA sequences of the
XX invention that encode for novel human secreted proteins.

XX Sequence 2037 BP; 367 A; 615 C; 565 G; 490 T; 0 other;

XX Query Match 3.2%; Score 60.4; DB 24; Length 2037;

XX Best Local Similarity 60.2%; Pred. No. 1.1e-05;

XX Matches 100; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 1132 TTGGGACCTCTGGAGACTGTGTGGCTTCGTGACGCCGGATCTGATGGAGCTGTT 1191
|||
Db 1235 TTCTCTACGCCAGCAGCTGTGCAAGCTTCTTCCCGGACATGAGAGGGGCTGCTC 1294

QY 1192 GTGTGGCCCTACTGCTGTGATCATGTTGTCGGGCTTGTGATGATGAGACATCAAC 1251
|||
Db 1295 ACCCTCCTGTTCATGCTTCTTCATCTTCACCTATGAGCCTGACATGATCTCAGCCTCAAG 1354

QY 1252 ACCATGACCCGCTTCGACATCCCAAGGCGCAAGACTATATCAATCA 1297
|||
Db 1355 ACCATGATGCTTGTGATGACCAAGAGGCCCCCATATATTTCTTGA 1400

RESULT 8

ABX71394
ID ABX71394 standard; cDNA; 2043 BP.

XX ABX71394;

XX 14-APR-2003 (first entry)

XX Human metabolism-associated cDNA from clone DKFzphes3_35b5.

XX Human; gene; gene therapy; vaccine; disease treatment; detection; ss.

XX Homo sapiens.

XX MO200112659-A2.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000MO-IB01496.

XX 18-AUG-1999; 99US-0149499.

XX 28-SEP-1999; 99US-0156503.

XX (GEMV-) GERMAN HUMAN GENOME PROJECT.

XX Wiemann S;

XX WPI; 2001-327840/34.

XX P-PADB; ABUS3209.

XX Nucleic acids having the sequences of clones isolated from libraries of
XX different human tissues, useful in recombinant DNA methodologies -

XX Claim 1; Page 830; 1095pp; English.

XX This invention describes novel polynucleotides and polypeptides isolated
XX from human cDNA libraries which can be used for gene therapy or in
XX vaccines. The polynucleotides of the invention and antibodies encoded by
XX them may be used in the prevention, diagnosis and treatment of diseases
XX associated with inappropriate polypeptide expression. The products of the
XX invention may also be used to identify modulators of expression and
XX activity and to down regulate expression and activity. The antibodies of
XX the invention may also be used as diagnostic agents for detecting the
XX presence of polypeptides in samples. This sequence encodes a polypeptide
XX described in the disclosure of the invention.

XX Sequence 2043 BP; 375 A; 617 C; 564 G; 487 T; 0 other;

XX Query Match 3.2%; Score 60.4; DB 23; Length 2043;

XX Best Local Similarity 60.2%; Pred. No. 1.1e-05;

XX Matches 100; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 1132 TTGGGACCTCTGGAGACTGTGTGGCTTCGTGACGCCGGATCTGATGGAGCTGTT 1191
|||
Db 1229 TTCTCTACGCCAGCAGCTGTGCAAGCTTCTTCCCGGACATGAGAGGGGCTGCTC 1288

QY 1192 GTGTGGCCCTACTGCTGTGATCATGTTGTCGGGCTTGTGATGATGAGACATCAAC 1251
|||
Db 1289 ACCCTCCTGTTCATGCTTCTTCATCTTCACCTATGAGCCTGACATGATCTCAGCCTCAAG 1348

QY 1252 ACCATGACCCGCTTCGACATCCCAAGGCGCAAGACTATATCAATCA 1297
|||
Db 1349 ACCATGATGCTTGTGATGACCAAGAGGCCCCCATATATTTCTTGA 1394

QY 1394 ACCATGATGCTTGTGATGACCAAGAGGCCCCCATATATTTCTTGA 1394

QY 1394 ACCATGATGCTTGTGATGACCAAGAGGCCCCCATATATTTCTTGA 1394

QY 1394 ACCATGATGCTTGTGATGACCAAGAGGCCCCCATATATTTCTTGA 1394

RESULT 9
AAK94606

ID AAK94606 standard; cDNA; 2063 BP.

```
XX AAK94606;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human full-length cDNA, SEQ ID NO: 3555.
DE
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1130094-A2.
PN
XX
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-0114089.
PF
XX
XX 08-JUL-1999; 99JP-0194486.
PR
XX
XX 11-JAN-2000; 2000JP-0118774.
PR
XX
XX 02-MAY-2000; 2000JP-0183765.
XX
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI
XX
XX Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
PI
XX
XX WPI; 2001-524255/58.
DR
XX
XX P-PSDB; AAM93671.
DR
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT
XX
XX use in genetic manipulation -
XX
XX
XX Claim 8; SEQ ID NO 3555; 1380bp + sequence listing; English.
PS
XX
XX The invention relates to primers for synthesizing full length cDNA
CC
XX
XX clones. 830 cDNA molecules encoding a human protein have been
CC
XX
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC
XX
XX molecules have been determined. Primers for synthesizing the full length
CC
XX
XX cDNA are useful for clarifying the function of the protein encoded by
CC
XX
XX the cDNA. The full length clones were obtained by construction of full
CC
XX
XX length enriched cDNA libraries that were synthesised by the oligo-capping
CC
XX
XX method. The primers enable the production of the full length cDNA easily
CC
XX
XX without any special methods. The present sequence is a full length
CC
XX
XX human cDNA of the invention.
CC
XX
XX Note: The sequence data for this patent did not form part of the printed
CC
XX
XX specification, but was obtained in CD-ROM format directly from EPO.
CC
XX
XX
XX
XX Sequence 2063 BP; 370 A; 622 C; 578 G; 493 T; 0 other;
SQ
XX
XX
XX Query Match 3.2%; Score 60.4; DB 22; Length 2063;
XX
XX Best Local Similarity 60.2%; Pred. No. 1.1e-05;
XX
XX Matches 100; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
XX
XX
XX 1132 TTGGGAGACTCTGGAGCTGTGGGCTTGTGACGCGCGGTATCCTGATGGAGCTGTT 1191
DB
XX
XX 1259 TTCTCTACGCCGACGACCTGTGCAAGCTTTCTCCCGCGCATGTGAGGGGCTGCTC 1318
XX
XX
XX 1192 GTGGTGCCCTTACTGCTGTGATCATGTTGTCGGGCTGTGCTGATGATGACATCAAC 1251
QY
XX
XX 1319 ACCTCCCTGTTCATGCTCTTCATCTTCACCTATGAGGCTTGACATGCTCAGGCTCAAG 1378
DB
XX
XX
XX 1252 ACGATGACCGCTTGCACGATCCCAAGGCGAAGATCATCA 1297
QY
XX
XX 1379 ACCATGATGCTTTGATGACCAAGGCGCCCACTATTTCTTGA 1424
DB
XX
XX
XX
XX RESULT 10
XX
XX AAX81298
XX
XX ID AAX81298 standard; cDNA; 2156 BP.
XX
XX AC AAX81298;
XX
XX DT 20-SEP-1999 (first entry)
```

```
XX
XX Human vacuolar ATPase subunit AC45 (HAC45) encoding cDNA.
DE
XX
XX Vacuolar ATPase subunit AC45; HAC45; vacuolar ATPase; ion transport;
XX
XX cancer; immune disorder; human; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX
XX FT CDS 111..1523
XX
XX FT /tag= a
XX
XX FT /product= "HAC45"
XX
XX
XX US5932444-A.
XX
XX
XX 03-AUG-1999.
PD
XX
XX 28-OCT-1997; 97US-0959011.
PF
XX
XX 28-OCT-1997; 97US-0959011.
PR
XX
XX 28-OCT-1997; 97US-0959011.
PR
XX
XX (INCY-) INCYTE PHARM INC.
PA
XX
XX Corley NC, Hillman JL, Shah P;
PI
XX
XX WPI; 1999-443597/37.
DR
XX
XX P-PSDB; AAY21800.
DR
XX
XX Vacuolar ATPase subunit AC45 for treating cancer and immune
PT
XX
XX disorders
XX
XX
XX Claim 4; Columns 37-40; 32pp; English.
PS
XX
XX This cDNA encodes a human vacuolar ATPase subunit AC45 (HAC45). Vacuolar
CC
XX
XX ATPase provides most of the energy required for ion transport processes
CC
XX
XX within the vacuolar system in eukaryotic cells. The polypeptide can be
CC
XX
XX produced recombinantly by culturing host cells transfected with a vector
CC
XX
XX comprising the HAC45 nucleic acid. Compositions comprising the HAC45
CC
XX
XX polynucleotide sequences are useful in the treatment of cancers and
CC
XX
XX immune disorders. The sequence complementary to the HAC45-encoding
CC
XX
XX sequence is useful for inhibiting expression of naturally occurring
CC
XX
XX HAC45.
CC
XX
XX
XX
XX Sequence 2156 BP; 388 A; 648 C; 605 G; 507 T; 8 other;
SQ
XX
XX
XX Query Match 3.2%; Score 60.4; DB 20; Length 2156;
XX
XX Best Local Similarity 60.2%; Pred. No. 1.2e-05;
XX
XX Matches 100; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
XX
XX
XX 1132 TTGGGAGACTCTGGAGCTGTGGGCTTGTGACGCGCGGTATCCTGATGGAGCTGTT 1191
DB
XX
XX 1344 TTCTCTACGCCGACGACCTGTGCAAGCTTTCTCCCGCGCATGTGAGGGGCTGCTC 1403
XX
XX
XX 1192 GTGGTGCCCTTACTGCTGTGATCATGTTGTCGGGCTGTGCTGATGATGACATCAAC 1251
QY
XX
XX 1404 ACCTCCCTGTTCATGCTCTTCATCTTCACCTATGAGGCTTGACATGATCTCAGGCTCAAG 1463
DB
XX
XX
XX 1252 ACGATGACCGCTTGCACGATCCCAAGGCGAAGATCATCA 1297
QY
XX
XX 1464 ACCATGATGCTTTGATGACCAAGGCGCCCACTATTTCTTGA 1509
DB
XX
XX
XX
XX RESULT 11
XX
XX ABR83610
XX
XX ID ABR83610 standard; cDNA; 2823 BP.
XX
XX AC ABR83610;
XX
XX DT 14-AUG-2002 (first entry)
XX
XX Human cDNA differentially expressed in granulocytic cells #181.
XX
XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
XX
XX
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM nucleic - nucleic search, using bw model

Run on: February 7, 2004, 18:37:05 ; Search time 116.011 Seconds
(without alignments)
7088.094 Million cell updates/sec

Title: US-09-614-150a-5

Perfect score: 1863
Sequence: 1 cgacgtcgagcaagcaagcaag.....agtttttcaatgtgttca 1863

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB. seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB. seq:*
3: /cgn2_6/ptodata/2/ina/5A COMB. seq:*
4: /cgn2_6/ptodata/2/ina/5B COMB. seq:*
5: /cgn2_6/ptodata/2/ina/PTUS COMB. seq:*
6: /cgn2_6/ptodata/2/ina/backfillseq1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.4	3.2	2156	2 US-08-959-011-2	Sequence 2, Appl1
2	48.4	2.6	1590	4 US-09-434-288-1	Sequence 1, Appl1
3	47.8	2.6	1584	4 US-09-252-991A-974	Sequence 974, App
4	47.8	2.6	1695	4 US-09-252-991A-1110	Sequence 1110, Ap
5	47.8	2.6	2766	4 US-09-252-991A-933	Sequence 933, App
6	47	2.5	2676	4 US-09-252-991A-12118	Sequence 12118, A
7	47	2.5	3618	4 US-09-252-991A-12305	Sequence 12305, A
8	45.4	2.4	909	4 US-09-252-991A-9153	Sequence 9153, Ap
9	45.4	2.4	1875	4 US-09-252-991A-8812	Sequence 8812, Ap
10	45.4	2.4	2517	4 US-09-252-991A-9238	Sequence 9238, Ap
11	43.8	2.4	7218	1 US-08-232-463-14	Sequence 14, Appl1
12	42.8	2.3	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
13	42.8	2.3	4411529	3 US-09-103-840A-1	Sequence 1, Appl1
14	42.6	2.3	804	4 US-09-773-748-2	Sequence 2, Appl1
15	42.2	2.3	1473	2 US-08-541-033A-25	Sequence 25, Appl1
16	42.2	2.3	1473	2 US-08-828-451-25	Sequence 25, Appl1
17	42.2	2.3	1506	2 US-08-541-033A-23	Sequence 23, Appl1
18	42.2	2.3	1506	2 US-08-828-451-23	Sequence 23, Appl1
19	42.2	2.3	1969	2 US-08-541-033A-7	Sequence 7, Appl1
20	42.2	2.3	1969	2 US-08-828-451-7	Sequence 7, Appl1
21	42.2	2.3	2096	2 US-08-541-033A-19	Sequence 19, Appl1
22	42.2	2.3	2096	2 US-08-828-451-19	Sequence 19, Appl1
23	42.2	2.3	2099	2 US-08-541-033A-3	Sequence 3, Appl1
24	42.2	2.3	2099	2 US-08-828-451-3	Sequence 3, Appl1
25	42.2	2.3	2137	2 US-08-541-033A-18	Sequence 18, Appl1
26	42.2	2.3	2137	2 US-08-828-451-18	Sequence 18, Appl1
27	42.2	2.3	2140	2 US-08-541-033A-1	Sequence 1, Appl1

28	42.2	2.3	2140	2 US-08-828-451-1	Sequence 1, Appl1
29	41.6	2.2	1608	4 US-09-252-991A-11665	Sequence 11665, A
30	41.6	2.2	1608	4 US-09-252-991A-11769	Sequence 11769, A
31	41.6	2.2	2304	4 US-09-252-991A-11691	Sequence 11691, A
32	41.2	2.2	2370	2 US-08-838-219B-19	Sequence 19, Appl1
33	41.2	2.2	2370	3 US-09-233-336A-19	Sequence 19, Appl1
34	41.2	2.2	2370	3 US-09-402-036-19	Sequence 19, Appl1
35	41.2	2.2	2370	3 US-09-233-336A-19	Sequence 19, Appl1
36	41.2	2.2	2370	3 US-09-402-036-19	Sequence 19, Appl1
37	41.2	2.2	2403	1 US-08-904-226-19	Sequence 19, Appl1
38	41.2	2.2	2403	1 US-08-471-033-30	Sequence 30, Appl1
39	41.2	2.2	2403	2 US-08-463-483A-30	Sequence 30, Appl1
40	41.2	2.2	2403	2 US-08-471-046A-30	Sequence 30, Appl1
41	41.2	2.2	2403	2 US-08-470-566B-30	Sequence 30, Appl1
42	41.2	2.2	2403	2 US-08-838-219B-7	Sequence 7, Appl1
43	41.2	2.2	2403	2 US-08-469-334-30	Sequence 30, Appl1
44	41.2	2.2	2403	3 US-09-300-529-30	Sequence 30, Appl1
45	41.2	2.2	2403	3 US-09-233-336A-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-08-959-011-2
; Sequence 2, Application US/08959011
; Patent No. 5932444
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: VACUOLAR ATPASE SUBUNIT AC45
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/959,011
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0412 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2156 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CRILNOT01
; CLONE: 676592
; US-08-959-011-2
Query Match 3.2%; Score 60.4; DB 2; Length 2156;
Best local Similarity 60.2%; Pred. No. 9.9e-07;

QY 946 AGCAATTTGGTCTACGACAAACAACCTTCGCTCCAGCGCGTGAATGCCCAACACG 1005
|||
DB 1602 AGCTACGTTGTTTACGCTGTGTGGCGGTGAGACTTCCTGCTGTTCGCGGGGTAC 1543
QY 1006 TTCTTTACTCTGCGGCACTTCACCTTGAAGTCGGCGCGTCAACAATGATAC 1065
1542 GTCTTCAACTACGATCGGCTGCGCATCGCGCGCGCGCTACGACCAAGGCGCG 1483
QY 1066 ACCCTGAGCTTCAAGTCCCTGACGCTGCAGCTCCCTTGAATGATACGACAGAG 1125
1482 CAAGCTGCTGATTCCTGCGCGGTGTGTGATCCTGCTGCTGCGGTCTTCAATAC 1423
DB 1126 TTTCCTTCCGCGCACTCTGCGGACTGTGTGGCTTCTGACGCGCGGTATCTGATG 1185
1422 GCCAATTTGGCGCTTCAGACACCTTCACGATCATCACTGTTCGCGACGCGCTTC 1363
QY 1186 CTGTTGTGTGCTGCGCTTACTGCTGTGATCATGTTCTGCGCGCTGCTGATGATG 1245
1362 GTGCTGACCAATCCCTGCTGCGCATCGCATCTCTTCAACCTTCGATGATCAGC 1303
QY 1246 ATCAACACGATGACCGCTTCAGACATCCCAAGGCGAAGTATCAACATCAATGCG 1305
1302 TACATCATGACGCTTACCGCGGATATCCCGCGACCAACCTGATGATCTTCCG 1243
DB 1306 GCC 1308
1242 GCC 1240

RESULT 5
US-09-252-991A-933
; Sequence 933, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 933
; LENGTH: 2766
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (2658)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-933

Query Match 2.6%; Score 47.8; DB 4; Length 2766;
Best Local Similarity 45.7%; Pred. No. 0.0037; Indels 0; Gaps 0;

Matches 166; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 946 AGCAATTTGGTCTACGACAAACAACCTTCGCTCCAGCGCGTGAATGCCCAACACG 1005
|||
DB 119 AGCTACGTTGTTTACGCTGTGTGGCGGTGAGACTTCCTGCTGTTCGCGGGGTAC 178
QY 1006 TTCTTTACTCTGCGGCACTTCACCTTGAAGTCGGCGCGTCAACAATGATAC 1065
1179 GTCTTCAACTACGATCGGCTGCGCATCGCGCGCGCGCTACGACCAAGGCGCG 238
DB 1066 ACCCTGAGCTTCAAGTCCCTGACGCTGCAGCTCCCTTGAATGATACGACAGAG 1125
239 CAAGCTGCTGATTCCTGCGCGGTGTGTGACCTTCTGCTGTGCGGTATCTTCAAGTAC 298
QY 1126 TTTCCTTCCGCGCACTCTGCGGACTGTGTGGCTTCTGATGACGCCGATATCTGATG 1185

DB 299 GCCAATTTGGCGCTTCAGACACCTTCACGATCATCACTTCGCTGCGATGACGCGCTTC 358
QY 1186 CTGTTGTGTGCTGCGCTTACTGCTGTGATCATGTTGTGTGGGTCTGCTGATGATGAC 1245
DB 359 GTGCTACCCCAATCTTGTGCTGCGCATCTCTTCAACCTTCGATGATCAGC 418
QY 1246 ATCAACACGATGACCGCTTTCAGATCCCAAGGCGAAGTATCAACATCAATGCGCG 1305
DB 419 TACATCATGACGCTTACCGCGGATATCCCGCGACCAACCTGATGATCTTCCG 478
QY 1306 GCC 1308
DB 479 GCC 481

RESULT 6
US-09-252-991A-12118
; Sequence 12118, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12118
; LENGTH: 2676
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12118

Query Match 2.5%; Score 47; DB 4; Length 2676;
Best Local Similarity 47.7%; Pred. No. 0.006; Indels 0; Gaps 0;

Matches 137; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 957 CTACGACAAACAACCTTCGCTCCAGCGCGTGAATGCCCAACACGTTCTTACTTC 1016
DB 831 CAAGGCTCTCAACATTTTCACTTCACTGCTGCTGCGCGCGCGCAACCGCACCGA 890
QY 1017 GTGCGCAACTCTACCTTCGATTCGCGCGCGCTCAACAATGATACACCTTGA 1076
DB 891 CCCTGCGCACAGCGGCTGAGGCGGTGTGCGCATGATTTCCAACTGATGCGG 950
QY 1077 CAAGTCCCTGACGCTGAGCTTCCTTCGATGATGATACAGAGAGATTTCTTCCG 1136
DB 951 CAACGGGCTCACTGCGCGGACTGTTCACATGTGCTCAAGAGAGGCTTCAACCTATT 1010
QY 1137 CGACTCTGGAAGTGTGCGGCTTCTGCGACGCGCGGTATCTGATGGAAGTGTGTG 1196
DB 1011 CCGCACTGGAAGTGTTCGCGCATGATACGCGGCTGCGGAGGCTGAGAGAGCT 1070
QY 1197 GCGCTTACTGCTGATCATGATTTGCTGCGGCTGCTGATGATGATG 1243
DB 1071 CCGCTTCTGCGCAACAACATGTTGCGCGAGAGCGCGGCGGATG 1117

RESULT 7
US-09-252-991A-12305/c
; Sequence 12305, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A

;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 12305
;; LENGTH: 3618
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12305

Query Match 2.5%; Score 47; DB 4; Length 3618;
Best Local Similarity 47.7%; Pred. No. 0.0072;
Matches 137; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 957 CTACGACAAACAACTTCCGCTCCAGCGCGTGAATGCCCAACCAAGTTCTTACTC 1016
DB 1873 CAAGGGGCTCAACATCTTCACTCCAGTTGCTGCGCCCGCCGAAACCGCACCGA 1814
QY 1017 GTGCGGCACTCACTCCGAGTCCGCGCGCGTCAACATGTAACAACACCTGAGCTT 1076
DB 1813 CGCGCGCCACACAGCGGCTCAAGGGCGTGGCGCATGATTTCCACAACAGTGGCG 1754
QY 1077 CAAGTCCCTGCACTGCAAGGCTCCCTTCATGATGTAAGTACAGAGAGACTTCCCTTCGG 1136
DB 1753 CAACCGGCTCACTGCGCGAGTGTTCGAAGTGTGCTCAAGAGAGGCTTCAACCGTAT 1694
QY 1137 CGACTCTGGAGCTGTGTGGGCTTCTGACGCCCGGATCTCTGATGGAGCTTTGTGGT 1196
DB 1693 CCAGGACTCGAGATTTCTCCCGCAATGAATCGCGGACAGCTCAAGCGCTGAGAGCGT 1634
QY 1197 CGCCCTACTGCTGATCATGTTGCTGCGGCGTCTGATGATG 1243
DB 1633 CGCCTTCTCTCGCACCAACGATTCGCCGAGAGCGCGGCGCATGG 1587

RESULT 8

US-09-252-991A-9153
;; Sequence 9153, Application US/09252991A
;; Patent No. 6551795
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 9153
;; LENGTH: 909
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9153

Query Match 2.4%; Score 45.4; DB 4; Length 909;
Best Local Similarity 47.4%; Pred. No. 0.0091;
Matches 136; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 929 GCGGATATTAGTCTGAGCAATTTGGTCTACGACAAACAACACTTCCGCTCCAGCGCG 988
DB 240 GCGGCGCATCAGATGGCATGAGACGGGTTGAACAACGCGATCTCCCGCGCACCGCA 239
QY 989 TGAATGCCCAACAGTTCTTACTGTCGGGCACTCAACCTCGAGTCCGCGCGCG 1048
DB 300 CGGCGGAGTACACCGGGAATTCACCGTGGCAACGAGGATGTCAACAGCGCG 359
QY 1049 TCAACAACATGTAACAACCGCTGAGCTTCAAGTCCCTGAGCTGACAGGCTCCCTTGATG 1108

DB 360 CGGTACCGCTATGCGACGACAGCTTGAGACCTTCGCGGACGTGATTCGCTGGCGCG 419
QY 1109 GTACGTACAAGAGACTTTCCTTCCGCGACTCTGGAGACTGTGTGGCTTGACGC 1168
DB 420 TCAAGTTCAAAAAGACAGACCGCGGACTATGCGGTGAAGAGTCCACACATGCGCG 479
QY 1169 CCGGATCTGATGGAGCTGTTGTGTGCTGCGCTTACTGCTGATC 1215
DB 480 CTTATGTGCTGCCGATGCCGAAAGGCGACGACATCAAGAGGTGCTC 526

RESULT 9

US-09-252-991A-8812/c
;; Sequence 8812, Application US/09252991A
;; Patent No. 6551795
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 8812
;; LENGTH: 1875
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8812

Query Match 2.4%; Score 45.4; DB 4; Length 1875;
Best Local Similarity 47.4%; Pred. No. 0.014;
Matches 136; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 929 GCGGATATTAGTCTGAGCAATTTGGTCTACGACAAACAACACTTCCGCTCCAGCGCG 988
DB 1679 GCGGCGCATCAGATGGCATGAGCGGTTGAACAACGCGATCTCCCGCGCACGCA 1620
QY 989 TGAATGCCCAACAGCTTCTTACTGTCGGGCACTCAACCTCGAGTCCGCGCGCG 1048
DB 1619 CGGCGGAGAGTACACCGCGGAATCACCGTGGCCACACAGCGCATCAACACAGCGG 1560
QY 1049 TCAACAACATGTAACAACCTGAGCTTCAAGTCCCTGAGCTGAGGCTCCCTTGATG 1108
DB 1559 CGGTCTACGCTATGCGACGACAGCTTGAAGACCTTGGGAGCTGATCGTGGGCG 1500
QY 1109 GTACGTACAAGAGACTTTCCTTCCGCGACTCTGGAGACTGTGTGGCTTGACGC 1168
DB 1499 TCAAGTTCAAAAAGACAGACCGCGGACTATGCGGTGAAGAGTCCACACATGCGCG 1440
QY 1169 CCGGATCTGATGGAGCTGTTGTGTGCTGCGCTTACTGCTGATC 1215
DB 1439 CTTATGTGCTGCCGATGCCGAAAGGCGACGACATCAAGAGGTGCTC 1393

RESULT 10

US-09-252-991A-9238
;; Sequence 9238, Application US/09252991A
;; Patent No. 6551795
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 9238
LENGTH: 2517
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9238

Query Match 2.4%; Score 45.4; DB 4; Length 2517;
Best Local Similarity 47.4%; Pred. No. 0.016;
Matches 156; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 929 GCGGATATTGCTGAGCAATTTGGTCTACGACAAACAATTCCGCTCAAGCGGCG 988
DB 929 GCGGCGCATGACATGGGATGAGCGGTTGAACAAGCCATGCTCCGCGCACGCCA 988
QY 989 TGAATGCCCAACAGCTTCTTACTGTCGCGCAACCTCACCTCGAGTCCGCGGCG 1048
DB 989 CGGCCGAGTACACCGGAAATCACCGTGGCAACGAGGATGTCACACAGCGCG 1048
QY 1049 TCAACAACATGTACACACCTGAGCTTCAAGTCCCTGACAGTGCAGGCTCCCTCGATG 1108
DB 1049 CGGCTACGCTTATGACAGCAAGCTTCAAGACCTGCGGAGCTGATCGCTGGGCG 1108
QY 1109 GTAGTACAAAGAGAGACTTCTCCCTGCGGACTCTGCGACTGTGTGGGCTTGTGAGCG 1168
DB 1109 TCAGGTTGAAAAGAGAGACCGGCGATGCGGTGAAGAAGTTCACACATGGGCG 1168
QY 1169 CCGGATCTGATGAGGACTGTTGTGTGTCGCGCTACGCTGCTGATC 1215
DB 1169 CTTATGCTGCCGATGCCGAAAGGACGACATCAAGAGTGTCTC 1215

RESULT 11

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT9Pdt-Fls
US-08-232-463-14

Query Match 2.4%; Score 43.8; DB 1; Length 7218;
Best Local Similarity 3.1%; Pred. No. 0.083;
Matches 12; Conservative 213; Mismatches 160; Indels 0; Gaps 0;

QY 651 GCAGTTCAGATGAGTGGCTTATCTGACTGCGCCGCCCTCAAGGCTCTGTGT 710
DB 1045 GCAGTTCAGAGGAGCTTGATGATTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1104
QY 711 GCAGCTGATCCCGCGGACACCGCCGACACCGGATGATGAGAGTCGAC 770
DB 1105 TT 1164
QY 771 CAACTGTTACATCTTCTACAGCTGCTCTACAGGCAACCCATCAGCTAC 830
DB 1165 TT 1224
QY 831 CAGCTCAAGCTCAACCACTAGCTCTACCAAGCTATCGTTATGACACATCTGT 890
DB 1225 TT 1284
QY 891 CGCCGACAAAGCAATACCTTGTGAGCTTTATTAATGGCGATTTAGTGTGAGCA 950
DB 1285 TT 1344
QY 951 TTTGCTACGACAAACAACACTCCGCTCCAGCGGAGTAAATCCCAACAGTCTTC 1010
DB 1345 TT 1404
QY 1011 TTAAGTGGGAGCACTCACTCTC 1035
DB 1405 TTTTTTTTTTTTTTTTTTTTTTTTTT 1429

RESULT 12

US-09-103-840A-2/C
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 2.3%; Score 42.8; DB 3; Length 4403765;
Best Local Similarity 45.9%; Pred. No. 5.8; 172; Indels 0; Gaps 0;
Matches 146; Conservative 0; Mismatches 172; Indels 0; Gaps 0;
QY 414 CGCCAGCTGACAGAGTGGAGCCCAAGCTTACTACACGAGCGTGAAGAACCTTCGGA 473

Db 1328748 CGGCCAGTTGCTCTCTGGCGCTGATGCGACCGTGTGTCAGCCTGTGTGAGACGCCCGCGCGG 1328689

QY 474 GGCACTGCGCTCACTGAGCGCCCAAGCGCGAGCACAACTCCATGATGCCAGTGGAAAGCT 533

Db 1328688 CTTGAGACTGCGCGGTGCGCTGCGACGCGGCTGATGATTCGACGACGACTGGCGGCTG 1328629

QY 534 GACCAAGCGCGCCCAAGTGTGCTGAGTACGCGCTGTTGTGACCTTTCAGAGACGCGCG 593

Db 1328628 GCGGCGCGCGCGCGGCTGCGCGCGATGTGTTCTTCTGATCGGGGTACCGACGACGAGCG 1328569

QY 594 CGAGAGCGCGGAGCGCGCTGAGTACATGACGCGCGCATAGCTCTCTATTCAGAGCA 653

Db 1328568 GGTGCGGTGGGTGACATGACGACGACGCGCGCGCGTGGCGCATCTTCTGCAAGCA 1328509

QY 654 GTTCGAGTGAAGGTGCTTATCTGTACTGCGCGCGCGCTCCAGCGCTCTGTTGTGCA 713

Db 1328508 CCGGTGCGACTCGGTGTGCGCGCGCGCGCGCTCGCGCGCTCGCGGTGTGAGCGGTGGA 1328449

QY 714 GCGTGTGACCGCGCGCTGA 731

Db 1328448 GCGCGCGCGCGCGCTGGA 1328431

RESULT 13

US-09-103-840A-1/c

Sequence 1, Application US/09103840A

Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103.840A

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 1

LENGTH: 4411529

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

OTHER INFORMATION: H37Rv

US-09-103-840A-1

Query Match 2.3%; Score 42.8; DB 3; Length 4411529;

Best Local Similarity 45.9%; Pred. No. 5.9;

Matches 146; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 414 CGCCAGCTGACGAGGAGTGAAGCCCAAGACCTTATACACGAGGTGAGAAACCCCTGGA 473

Db 1329280 CGGCACTTCTCTCTGCGCTGATGCGACCGTGTGACGCTGTGTGAGCGCCCGCGG 1329221

QY 474 GGCACTGCGCTCACTGAGCGCCCAAGCGCGAGCACAACTCCATGATGCCAGTGGAAAGCT 533

Db 1329220 CTTGAGACTGCGCGGTGCGCTGACGCGGCTGATGATTCGACGACGACTGGCGGCT 1329161

QY 534 GACCAAGCGCGCCCAAGTGTGCTGAGTACGCGCTGTTGTGACCTTTCAGAGACGCGCG 593

Db 1329160 GCGGCGCGCGCGCGGCTGCGCGCGATGTGTTCTTCTGATCGGGGTACCGACGACGAGG 1329101

QY 594 CGAGAGCGCGGAGCGCGCTGAGTACATGACGCGCGCATAGCTCTCTATTCAGAGCA 653

Db 1329100 GGTGCGGTGGGTGACATGACGACGACGCGCGCGCTGTGCGCATCTTCTGCAAGCA 1329041

QY 654 GTTCGAGTGAAGGTGCTTATCTGTACTGCGCGCGCGCTCCAGCGCTCTGTTGTGCA 713

Db 1329040 CCGGTGCGACTCGGTGTGCGCGCGCGCGCTCGCGCGCTCGCGGTGTGAGCGGTGGA 1328981

QY 714 GCGTGTGACCGCGCGCTGA 731

Db 1328980 GCGCGCGCGCGCGCTGGA 1328963

RESULT 14

US-09-773-748-2

Sequence 2, Application US/09773748

Patent No. 6534297

GENERAL INFORMATION:

APPLICANT: Wada, Masaru

TITLE OF INVENTION: Levodione reductase gene and use thereof

FILE REFERENCE: Levodione reductase

CURRENT APPLICATION NUMBER: US/09/773,748

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: EP00101665.8

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 2

LENGTH: 804

TYPE: DNA

ORGANISM: Corynebacterium aquaticum

US-09-773-748-2

Query Match 2.3%; Score 42.6; DB 4; Length 804;

Best Local Similarity 44.6%; Pred. No. 0.051;

Matches 168; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 955 GTCTAGACAAACAACACTTCCGCTCCAGCGCGGTGATGCCCCCAACGATTTCTTTAC 1014

Db 190 GCGAGGTCTCTACCAACGCTCGCGCGATGCTGGAGAGAGGCCAGGCTGAGGCTACGTC 249

QY 1015 TCGTGGGCAACTCACTCGAGTCCGCGCGCTGCAACAACATGTAACAACACCTGAGC 1074

Db 250 ACCGCACACACGAGGCTTCGCGCGCATGACGCGCTTCTTCAACAACGCGCGCATCGAG 309

QY 1075 TTGAATCCTGTGAGCTGACGAGGCTCCCTTGATGATGATGATGATGATGATGATGATGAT 1134

Db 310 GCGAAGAGAAACCGACCGAGTGTTCACGCGCGCGGATTCGACAAAGTCTGTGATC 369

QY 1135 GCGACTCTCTGAGCTGTGTGAGCTTCTGACGCGCGGTATCTGATGAGGATGTTGTG 1194

Db 370 AACCTGCGCGCGCTGTCTCTCGCGCTCGAGAGTCTGAAATGATGCGGACGAGGCG 429

QY 1195 GTGCGCTTCTGCTGATCATGTTGTGCGCGCTTCTGCTGATGATGATGATGATGATGATGAT 1254

Db 430 TCCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489

QY 1255 ATGACCGCTTTCAGCATTCGACGCGGACGACGATCATGATCATGATCATGATCATGATCAT 1314

Db 490 GGTACGCGCGCGCGCAAGCAAGGAGTGTGCTGATGATGATGATGATGATGATGATGAT 549

QY 1315 ACCGCTTACATCGG 1331

Db 550 GCGCGCTACGATCGG 566

RESULT 15

US-08-541-033A-25

Sequence 25, Application US/08541033A

Patent No. 5879941

GENERAL INFORMATION:

APPLICANT: Schmidt, Robert R.

APPLICANT: Miller, Philip

TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES

TITLE OF INVENTION: RELATING TO THE "- AND -SUBUNITS OF GLUTAMATE

TITLE OF INVENTION: DEHYDROGENASES AND METHODS OF USE

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanichik & Saliwanichik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/541,033A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UP155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ. ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1473 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 4..1431
US-08-541-033A-25

Query Match 2.3%; Score 42.2; DB 2; Length 1473;
Best Local Similarity 47.5%; Pred. No. 0.093;
Matches 125; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY	1023	CAACCTCACCTCGAGTCGCGGCGGCTCAACACATGATACACACCGCTTCAAGTC	1082
DB	237	CAAGCAGATGCTTGAAGCTGAGCGCGTATCACTTCGGGTCTGGCTGACGACGC	296
QY	1083	CCTGACGCTGACAGGCTCCCTTCATGATGATACGTACAGAGAGACTTCCCTTCGCGCATC	1142
DB	297	CGGCAACTGACAGGTCAACCGCGGCTTCGCGTCACTACTCGTCCGCCATCGGCCCTTA	356
QY	1143	CTGGGACTGTGTGGGCTTCTGTACGCGCCGATATCTGATGAGACTGTTGTGTGCGCCT	1202
DB	357	CAAGGCGGCTGCGCTTCCACCCCTCCGGAACCTGTCCATCATGAAAGTTCTTGCCTT	416
QY	1203	ACTGCTGATCATATGTTGTCGCGGCTTCTGATGATGATGATGATGATGATGATGATG	1262
DB	417	TGAGCAGATCTTAAAGAACAGCTTGAACACCTTCCCATGCGCGCGCAAGGCGGCTC	476
QY	1263	CTTGACGATCCCAAGGCGCAGA	1285
DB	477	CGACTTCGACCCCAAGGCGCAGA	499

Search completed: February 8, 2004, 10:31:18
Job time: 129.011 secs

THIS PAGE BLANK (USPTO)


```
RESULT 5
US-10-017-161-2047/c
; Sequence 2047, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2047
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1231)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1031)
US-10-017-161-2047

Query Match      2.7%; Score 51; DB 13; Length 1231;
Best Local Similarity 46.7%; Pred. No. 0.00025;
Matches 162; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY      732 CACCGCGCCACAACCGGTGGCATGATGGAAGTGCACCAATCAGTTTGAATCTTCTA 791
      |||||
DB      872 CACCGCACCATCACCCTCGTCACCATTAATACACACCAACCATCACTACCAACAC 813

QY      792 CACTGCGCTGCTTACAAAGGCAACCCCATACCGGTACCGAGCTCAAGCTCAACCAATC 851
      |||||
DB      812 CACTTTACCAACCAACCAACCAACCAACCATCCTCGTCACTATATACCAACCAAC 753

QY      852 TAGCTTACCAAGCTATCCGTTGTTATGACATATGTCGCGCAAGCAAGCATCACTT 911
      |||||
DB      752 CATCATTAACCAACCAACCACTTCACACCAACCAACCAACCAATCTTCTCATACCAT 693

QY      912 TGAAGTTGTTATATGAGCGGATATCTTACTGAGCAATTTGGTCTACGACAACAA 971
      |||||
DB      692 CACCATTAACAAGCTGTCGATAGCCATGACATGACACCAACCATCACTACTATATAC 633

QY      972 CTTCGCTCCAGCGGCGGTGAATGCCCAACCAACGTTCTTACTCGTGGGCAACCTGAC 1031
      |||||
DB      632 TGCCACCAACCATCAACCAACCAACCAACCATCACTTCACCAACCAACCTTCATAC 573

QY      1032 CCTCGAGTCGCGCGCTCAACAACATGTACAACCCCTAGCTTCA 1078
      |||||
DB      572 CATCAACATCACTACCAACCAACCAACCACTTCACCAACATACCAACCA 526

RESULT 6
US-10-292-798-1629
; Sequence 1629, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT FILING DATE: 2002-11-13
```

```
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1629
; LENGTH: 1972
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(1972)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1772)
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (975)..(594)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (1319)..(1328)
; OTHER INFORMATION: a, t, c, g, unknown or other
US-10-292-798-1629

Query Match      2.7%; Score 51; DB 12; Length 1972;
Best Local Similarity 46.7%; Pred. No. 0.00034;
Matches 162; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY      732 CACCGCGCCACAACCGGTGGCATGATGGAAGTGCACCAATCAGTTTGAATCTTCTA 791
      |||||
DB      266 CACAGCACCATCACCATATATACCAACCAACCAACCATATACCAACCAATAC 325

QY      792 CACTGCGCTGCTTACAAAGGCAACCCCATACCGGTACCGAGCTCAAGCTCAACCAATC 851
      |||||
DB      326 CGGCACCAACCATCAACCATATATATATACAGGCAACCAACCAACCATATATACCAAC 385

QY      852 TAGCTTACCAAGCTATCCGTTGTTATGACATATGTCGCGCAAGCAAGCATCACTT 911
      |||||
DB      386 CACCAACCAACCATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 445

QY      912 TGAAGTTGTTATATGAGCGGATATCTTACTGAGCAATTTGGTCTACGACAACAA 971
      |||||
DB      446 CATCATATATATACAGCCACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 505

QY      972 CTTCGCTCCAGCGGCGGTGAATGCCCAACCAACGTTCTTACTCGTGGGCAACCTGAC 1031
      |||||
DB      506 CACCATCAACATACCAACCATACCGGCAACCAACCAACCAACCAACCAACCAACCAAT 565

QY      1032 CCTCGAGTCGCGCGCTCAACAACATGTACAACCCCTAGCTTCA 1078
      |||||
DB      566 CATCAACATCAACCATATATACCAACCAACCAACCAACCAACCAACCAACCAAT 612

RESULT 7
US-10-017-161-1981
; Sequence 1981, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
```

```

; SEQ ID NO 1981
; LENGTH: 1972
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1972)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1772)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (975)..(994)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1319)..(1328)
; OTHER INFORMATION: a, t, c, g, unknown or other
US-10-017-161-1981
```

```

Query Match          2.7%; Score 51; DB 13; Length 1972;
Best Local Similarity 46.7%; Pred. No. 0.00034;
Matches 162; Conservative 0; Mismatches 185; Indels 0; Gaps 0;
```

```

QY 732 CACCGCGCCCAACACCGGTGGCATCATGTGGAAGTCGACCAATCATGTTTCAATCTTCTA 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 266 CACAGCCACCAACCAATCATCATCAACCAACCAACCAACCAACCAATCATCATCAAC 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 792 CACTGCCCTGCTTACACGGCAACCCCATCATCCGTACCGGACTTCAAGCTCAACCAATC 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 326 CGCACCACCATCATCAATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 852 TAGCTTACCAAGCTATCCGTTGTTATGACATCTGTGGCGCAACCAAGCAATCACTT 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 386 CACACACCAACCAATCATCAACCAATCATCATCATCATCATCATCATCATCATCATCAT 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 912 TGAAGTTGTTTATATGCGGATATCTTATGTCGAGCAATTTGGTCTACGACCAACAA 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 446 CATCATCATATATACGACGCAACCAACCAATCATCATCATCATCATCATCATCATCATCA 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 972 CTTCGGCTCCAGGGCGGTGAATGCCCCCAACCAAGTTCTTTACTGTCGGGCAACTTCAC 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 506 CACCATCATCATCATCAACCAATCATCAACCGGACCATCATCATCATCATCATCATCATCA 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1032 CCTCGAGTCGGCGCGGTCAACAACATGTACAACACCTGAGCTTCA 1078
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 566 CATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 8
US-10-029-386-25133
; Sequence 25133, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25133
; LENGTH: 1493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP001751.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
```

```
US-10-029-386-25133
```

```

Query Match          2.7%; Score 49.4; DB 13; Length 1493;
Best Local Similarity 46.4%; Pred. No. 0.00087;
Matches 161; Conservative 0; Mismatches 186; Indels 0; Gaps 0;
```

```

QY 732 CACCGCGCCCAACACCGGTGGCATCATGTGGAAGTCGACCAATCATGTTTCAATCTTCTA 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 CACCATCATCAACCAATCATCAACCAATCATCATCATCATCATCATCATCATCATCATCA 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 792 CACTGCCCTGCTTACACGGCAACCCCATCATCCGTACCGGACTTCAAGCTCAACCAATC 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 249 CACCATCATCAACCAATCATCAACCAATCATCATCATCATCATCATCATCATCATCATCA 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 852 TAGCTTACCAAGCTATCCGTTGTTATGACATCTGTGCGCAACCAAGCAATCACTT 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 309 CAACACCAACCAATCATCAACCAATCATCAACCAATCATCATCATCATCATCATCATCATCA 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 912 TGAAGTTGTTTATATGCGGATATCTTATGTCGAGCAATTTGGTCTACGACCAACAA 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 369 CATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 972 CTTCGGCTCCAGGGCGGTGAATGCCCCCAACCAAGTTCTTTACTGTCGGGCAACTTCAC 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 429 CACCATCATCATCATCAACCAATCATCAACCAATCATCATCATCATCATCATCATCATCA 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1032 CCTCGAGTCGGCGCGGTCAACAACATGTACAACACCTGAGCTTCA 1078
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 489 CATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 9

```

US-10-271-889-1/c
; Sequence 1, Application US/10271889
; Publication No. US20030194784A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA Encoding Methymycin and Pikromycin
; FILE REFERENCE: 600,582051
; CURRENT APPLICATION NUMBER: US/10/271,889
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 09/861,289
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 09/860,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 09/836,821
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Streptomyces narbonensis
US-10-271-889-1
```

```

Query Match          2.6%; Score 48.4; DB 13; Length 1590;
Best Local Similarity 46.9%; Pred. No. 0.0018;
Matches 151; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
```

```

QY 516 CGATGCCAGTGAAGCTGACACGCGGCAAGTGTGCTGAGTACGGCCCTGTTTGT 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1248 CGACGGCATCCGCGCGGATCATCCCGTGTGACTACCTCCACAGCCCAACGTCGA 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 576 GACCTTCGAGGAGCGCGCGAGAGCGGAGCGAGCTGAGATCATGACGCGGCAT 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1188 GACCATTCGAGAGAGCTGCGACGCTCTGCGCGGACTACCCCGGACCTCCGCGCT 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 636 AGCTGTATCAGCAGAGTTGAGTGCAGGTGCTTATCTGTACTGTGCGCCGCCCTTC 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 1128 CCCCTTCTTCTCCACCGTCGACCGTGCCTGGATCACCAGAACCGCCCTCGACCCACGTA 1069
Qy 696 CACGCGCTCTGTGTGAGCGTGTATCCGCGCGTGAACCGCGCCGACCAACGGGTGGCAT 755
Db 1068 CTGGTACCCGACACTCCGCCACCCCGTGGCTTGGCCCGCCGCGTGAACCTCTCGCCAC 1009
Qy 756 CATGTGAAGTCGACCAATCAGTTTCAATCTTCACTGCGCTGCTTACACGCGCA 815
Db 1008 CCAAGCAAGGTTTACCCACTTGTGTCAGGTGACGGCCACCCGCTCTCAACATGGCCCT 949
Qy 816 CCCCATCACCCTGACCGGACTC 837
Db 948 CCCGACACCGTCACCGGCTC 927

RESULT 10
US-10-156-761-6160
; Sequence 6160, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 6160
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1371)
US-10-156-761-6160

Query Match 2.6%; Score 48.2; DB 15; Length 1371;
Best Local Similarity 42.5%; Pred. No. 0.0019;
Matches 315; Conservative 0; Mismatches 423; Indels 3; Gaps 1;

Qy 568 CTGTTTGTACCTTTCAGAGACCGCCGCGAGAGCGCGGAGCGCTGTGAGTCAATGAC 627
Db 184 CTGGCCGAGACCTGTGCGCGAATGCGCGGACTTCCGAGCGCGGAGCAAGAG 243
Qy 628 GCCGCAATAGCTCTATCAAGCAAGATTGAGTGAAGTGTGCTTATCTGATCGCC 687
Db 244 CCGACGAGCTGCGCAATCAAGGCGCCGCTGCGCGAAGTGGCGCTGCGGCCACAGCC 303
Qy 688 GCCCCTTCAAGCGCTCTGTGTGTCAGCGTGTACCCCGCGGTGACACCGCGCACAC 747
Db 304 TCCCGCGCACGCGCACCGCGCCACGACCGCTGGAAGCGCGCTGGGCAAGAGGCC 363
Qy 748 GGTGATCATGTGGAAGTGAACCAATCAATCTTCAATCTTCACTGCGCTGTAC 807
Db 364 GGTGCTCTCTGGGCAAGCGGTGAGAGACTCCCTTCAGCGCATCGCGCGCTCGCC 423
Qy 808 AACGGCAACCCCACTCAACCGTCAACGCTCAAGCTCAACCACTTACCAAGCTA 867
Db 424 CCGGCAACCGGCACTGCGCCCTGCAACCTGTGTTACAGGCGCGCGCTCCCGCGAA 483
Qy 868 TCGGTGTATGAGCAATCTGTGCGCGCAACCAACCAATCACTTTGAGCTTTTAAAT 927
Db 484 CTGGCGGCGGCGCACCGCTTGAACCGCGCGCTGCGCAACCACTTCCCTCGCGAAGCATC 543

Qy 928 GGGGATATCTTATAGTCTGAGCAATTTGTCTACACAAACAATTCGCTTCAGCGGC 987
Db 544 GACGGATCCCGGAGAGACAGACCTCAACTACCCCTCTGTGAATCTGTGTGCTCAG 603
Qy 988 GTGAATGCCCAACCAAGCTTCTTATGCTGTGGGCAACTTCACTTGAATCGCGGCC 1047
Db 604 CCGCAAGGCAAGGCTTTCAGCAACCGCGGAAGTCCCGCTTGTGCTTCAAGCAATCCCC 663
Qy 1048 GTCAACAATGTATCAACAACCGTGAAGTTCAGTCCCTGAGGTGAGGCTCCCTTGAT 1107
Db 664 GCGGCGGCACTTACCGCGAGGCAATGCGTACCGCAACTCTTCAACGCAATCGAG 723
Qy 1108 GTTACGTACAGAGAGACTTTCCTTGGCGACTCTTGGAGCTGTGTGAGCTTCTGAGC 1167
Db 724 CCGCGCGCACGG---CCGCGCAACGCAACCGGTTCCGGAAGTGAATGGCGGCTGATC 780
Qy 1168 CCGGTATCTGTATGGAAGTCTGTTGTGTGCGCCCTACTGCTGTCAATGTTCTGCGC 1227
Db 781 CCGGCGAGCTGACCGCTGAGCAACCGGAGAACCCGCGCGCCCGCGCCAGGCC 840
Qy 1228 GTCTGTGATGATGAGCAATCAACAGATGACCGCTTGAAGATCCAGGCAAGGCAACT 1287
Db 841 CACCGGACGCCACCTTACCCACCGCAACGCGGAGCGGTCTACGCGGATGTTCAAGGCC 900
Qy 1288 ATCACCATCAATGCGCGGCC 1308
Db 901 GCCACCATGCGCGAGGCCGCC 921

RESULT 11
US-10-292-798-1695/c
; Sequence 1695, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 08435/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1695
; LENGTH: 38918
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(38918)
; NAME/KEY: CDS
; LOCATION: (201)..(228)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8843)..(8927)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19620)..(20268)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25302)..(25419)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26738)..(26852)
; FEATURE:

NAME/KEY: CDS
LOCATION: (31546)..(31644)
FEATURE:
NAME/KEY: CDS
LOCATION: (38460)..(38718)
US-10-792-798-1695

Query Match
Best Local Similarity 46.1%; Pred. No. 0.017;
Matches 161; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

735 CGCGCCCAACACCGGTGGCATCATGTGGAAGTGGACCAATCATGTTCAATCTTTACAC 794
20280 CACCAACATCATCATCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 20221
795 TGCCCTGCTCTACACGGGCAACCCCATCATCATCATCATCATCATCATCATCATCATCATCAT 854
20220 CCCCACTGTGTGTCACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 20161
855 CTCTACCAAGCTATCCGTTGTTATGACACATCTGTGCGGCAAGCCAACTCACTTTGA 914
20160 CATCAACATCATCATCAACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 20101
915 CGTTGTTATTAATGCGGATATCTTGTAGTGAAGCAATTTGCTTACGACAACAACAACTT 974
20100 CACTATCAACCACTCAACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 20041
975 CGCCTCAAGGGGGTGAATGCCCAACCAAGCTTCTTTACTCGGCGGCAACCTCAACCTT 1034
20040 CACCAACATCATCATCAACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 19981
1035 CGAGTCGGGGCGGTCAACAACATGTACAAACCAACCTGAGCTTCAAGTCC 1083
19980 CATCAACCTCATCATCAACCTTCAACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 19932

RESULT 12
US-10-017-161-2049/c

Sequence 2049, Application US/10017161
Publication No. US20030143668A1

GENERAL INFORMATION:

APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABEURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS

FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18

PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18

NUMBER OF SEQ ID NOS: 2430
SOFTWARE: Patent Ver. 2.1

SEQ ID NO 2049

LENGTH: 38918

TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: SOURCE
LOCATION: (1)..(38918)

FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(228)

FEATURE:
NAME/KEY: CDS
LOCATION: (8843)..(8927)

FEATURE:
NAME/KEY: CDS
LOCATION: (19620)..(20268)

FEATURE:
NAME/KEY: CDS
LOCATION: (25302)..(25419)

FEATURE:

FEATURE:

FEATURE:

NAME/KEY: CDS
LOCATION: (26738)..(26852)
FEATURE:
NAME/KEY: CDS
LOCATION: (31546)..(31644)
FEATURE:
NAME/KEY: CDS
LOCATION: (38460)..(38718)
US-10-017-161-2049

Query Match
Best Local Similarity 46.1%; Pred. No. 0.017;
Matches 161; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

735 CGCGCCCAACACCGGTGGCATCATGTGGAAGTGGACCAATCATGTTCAATCTTTACAC 794
20280 CACCAACATCATCATCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 20221
795 TGCCCTGCTCTACACGGGCAACCCCATCATCATCATCATCATCATCATCATCATCATCATCAT 854
20220 CCCCACTGTGTGTCACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 20161
855 CTCTACCAAGCTATCCGTTGTTATGACACATCTGTGCGGCAAGCCAACTCACTTTGA 914
20160 CATCAACATCATCATCAACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 20101
915 CGTTGTTATTAATGCGGATATCTTGTAGTGAAGCAATTTGCTTACGACAACAACAACTT 974
20100 CACTATCAACCACTCAACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 20041
975 CGCCTCAAGGGGGTGAATGCCCAACCAAGCTTCTTTACTCGGCGGCAACCTCAACCTT 1034
20040 CACCAACATCATCATCAACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 19981
1035 CGAGTCGGGGCGGTCAACAACATGTACAAACCAACCTGAGCTTCAAGTCC 1083
19980 CATCAACCTCATCATCAACCTTCAACCAACCAACCAACCAATCATCATCATCATCATCATCATCATCATCAT 19932

RESULT 13

US-10-156-761-1/c

Sequence 1, Application US/10156761
Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, YUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 1

LENGTH: 9025608

TYPE: DNA

ORGANISM: Streptomyces avermitilis

FEATURE:

NAME/KEY: misc feature

LOCATION: (418715)

OTHER INFORMATION: a, t, c, g, other or unknown

US-10-156-761-1

Query Match

Best Local Similarity 42.5%; Pred. No. 0.61;

Matches 315; Conservative 0; Mismatches 423; Indels 3; Gaps 1;

```
QY 568 CTGTTGTGACCTTTCAGAGAGCGCCCGAGAGCCCGAGGCGCATCTGAGTCAATGAC 627
DB 7430121 CTGGCCGAGGACCTGCTGGCGCAATGCGGACCTGCGGAGCGGCTGGCGGACGAGAG 7430062
QY 628 GCGGCATAGCTGCTATCAGAGAGAGAGTTCAGTGAAGTGGCTTATCTGTACTGTCGGC 687
DB 7430061 CCGACGAGACTGGCGAACAATCAAGGCGCCCGTGGCCGAGTGGCGCTGCGGCGCCAGGCC 7430002
QY 688 GCGCCCTCCAGCGGCTCTGTGTGAGAGCGTGTACCGCGCGTACACCGCGCGCAACGC 747
DB 7430001 TCCCCCGGACCGCACCACCGCGGACGAGCGTGTGAAGCGCGCTGGCTGGGAGAGCGGCC 7429942
QY 748 GGTGCGATCATGTGAAGTGCACCAATCACTTTTCAATCTTCTTACAATGCTGCTTAC 807
DB 7429941 GGTGCGCTCTGTGGCAAGCGGCTGAGAGAGCTCCCGCTGAGCGGACATCGCGCGCTCGCC 7429882
QY 808 AACGCGAACCCCATCAACCGTCAACCGACCTCAAGCTCAACCACTCTAGCTTACCAAGCTA 867
DB 7429881 CCGGCGACCGGCAACTGGCGCCCTGACACCTGTGTTACGCGCCGCGCGCTCCCGCGGAA 7429822
QY 868 TCGGTTGTTATGAGCAATCTGTGTGCGGACCAAGCGCAATCACTTGAAGTTGTTATTAAT 927
DB 7429821 CTGGCCGCGGCGCAACCTGTGAACCGCGCTGCGCAACCACTCTCGCGCGAGACATC 7429762
QY 928 GCGGATTAATTAATGCTGAGCAATTTGTCTTACAGCAACAACCTTCCGCTTCAAGCGGC 987
DB 7429761 GACGCGATCGCGGAGAGACGACGACTCAACTACCCCTCTGTAACATCTCTGTGCTCAG 7429702
QY 988 GTGAATGCCCAACAACGCTTCTTCTTACTGTGTGCGGCAACCTCACTGAGTCCGCGGCC 1047
DB 7429701 CCGCACGCGGAGAGGCTTTCAGACCGCGGAGTGGCGCGCTGCTGCTGACGAACTCCCC 7429642
QY 1048 GTCAACAATGATGACAACCCGAGCTTCAAGTCCCTGAGTGTGAGGCTCCCTTCGAT 1107
DB 7429641 GCGGCGCGACGTTACCGCGGAGCGCATGCGTACCGCAACCTCTGACGCGCATCGAG 7429582
QY 1108 GGTACGTACAAGAGAGACTTTTCCCTTGGCGACTCTCTGAGACTGTGTGGGCTTCTGTGAG 1167
DB 7429581 CCGCGCGGACGCG---CCGCGCACCGGCAACCCCTTCCGAGAGTGAATCGCGCGCTGATC 7429525
QY 1168 CCGGATTCCTGATGAGAGCTTTTGTGTGCGCCCTTACTCTGTGTCATATGTTGTGTGCGC 1227
DB 7429524 CCGCGCGAGGTGACGCGCTGACCAACCGGAGAGACCGCGGCGCGCGCGCGCGCGCGCC 7429465
QY 1228 GTCTGTGATGATGATGACATCAACAGATGAGACCGCTTGAAGATCCCAAGGCGAAGT 1287
DB 7429464 CACCGCGAGCGCAACCTCTCAACCAACGCGGAGCGGCTTACGCGGATGTTCAAGGCC 7429405
QY 1288 ATCACCATCATGCGCGCGCC 1308
DB 7429404 GCCACCATGCGCGAGCGCGCC 7429384

RESULT 14
US-10-228-063-4
; Sequence 4, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Iannan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2400
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-228-063-4
```

```
Query Match 2.6%; Score 47.8; DB 13; Length 2400;
Best Local Similarity 45.7%; Pred. No. 0.0037;
Matches 166; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 956 TCTAGACAACAACAACCTTCCGCTCAGCGGGGTGATGATGCCCAACCAACGTTCTTACT 1015
DB 1979 TCTGCGGACCAACAACCTTCAACGACACCTCTTACACCGCCCGATCTTCATCAACGCGCT 2038
QY 1016 CGTGGGCAACCTCAACCCCTGAGTCGCGGCGGCGTGAACAACATGTAACAACCTGAGGT 1075
DB 2039 TCGACTACGAGCGAAGCTTCAGTTTCATGAGTGTTCATCAACCAAGGCGCTCATCA 2098
QY 1076 TCAAGTCCCTGAGCTGAGGCTCCCTTTCATGATGATGATGATGATGATGATGATGATGAT 1135
DB 2099 AGCTCCGAGAGAGACACCGCGCTTCCGCTCAAGAACCCGAGAGATCAAGAACAC 2158
QY 1136 GCGACTCTGGAGACTGTGTGGCTTGTGACGCGCGGTATCTGTATGGAGCTGTTGTGG 1195
DB 2159 TGAAGTCTCCCGGCGGCGCGCGCATGTGGCTTCAATGCTCAAGGACACGCGCGCG 2218
QY 1196 TCGCCCTACTGCTGTGATCATATGTTCTGTGGGCTGTGCTGATGATGATGATGATGATGAT 1255
DB 2219 GCGACCTCGTGAAGACATGTGTGATCTCAACGCGCAACCTGAGAGAGACCACTTCA 2278
QY 1256 TGAACCGCTTCAACGATCCCAAGGCGAAGATATCAACCATCATGATGCGCGCGAGTAA 1315
DB 2279 AGCTCCGAGAGGCAAGTGAACGTGTGTGAATCTCCAGAAAGCGCGGACCGAGGTGA 2338
QY 1316 CCG 1318
DB 2339 TCG 2341
```

```
RESULT 15
US-10-228-063-25
; Sequence 25, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Iannan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 2478
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-228-063-25

Query Match 2.6%; Score 47.8; DB 13; Length 2478;
Best Local Similarity 45.7%; Pred. No. 0.0037;
Matches 166; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 956 TCTAGACAACAACAACCTTCCGCTCAGCGGGGTGATGATGCCCAACCAACGTTCTTACT 1015
DB 2036 TCTGCGGACCAACAACCTTCAACGACACCTCTTACAGCGCCCGATCTTCATCAACGCGCT 2095
QY 1016 CGTGGGCAACCTCAACCCCTGAGTCGCGGCGGCGTGAACAACATGTAACAACCTGAGGT 1075
DB 2096 TCGACTACGAGCGAAGCTTCAGTTTCATGAGTGTTCATCAACCAAGGCGCTCATCA 2155
QY 1076 TCAAGTCCCTGAGCTGAGGCTCCCTTTCATGATGATGATGATGATGATGATGATGATGAT 1135
DB 2156 AGCTCCGAGAGAGACACCGCGCTTCCGCTCAAGAACCCGAGAGATCAAGAACAC 2215
QY 1136 GCGACTCTGGAGACTGTGTGGCTTGTGACGCGCGGTATCTGTATGGAGCTGTTGTGG 1195
DB 2216 TGAAGTCTCCCGGCGGCGCGCATGTGGCTTCAATGCTCAAGGACACGCGCGCG 2275
```

Qy	1196	TCGCCCCCTACTGCTGTCATCATGTCGCGCGCTGCTGATGATGACATCAACACGA	1255
Db	2276	GCGACCCCGTGAAGAGACATGCTGTGATCTACAAAGGCACTGAGAAAGACCACTACA	2335
Qy	1256	TGGACCGCTTCGACGATCCCAAGGGCAAGCTATCAACCATCAATGCCGCCCGAGTAA	1315
Db	2336	AGCTCCCGAGGCGCAAGTGAACGTGGTGTGAATCCCAAGAGCCGCGCACCGAGGTGA	2395
Qy	1316	CCG	1318
Db	2396	TCG	2398

Search completed: February 8, 2004, 20:44:29
 Job time : 720.824 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:36:25 ; Search time 4369.39 Seconds
(without alignments)
10362.824 Million cell updates/sec

Title: US-09-614-150A-5

Sequence: 1 gagctcgagcaagcaag.....agtttttcaattgtgtta 1863

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_ges_hum: *
18: em_ges_inv: *
19: em_ges_pin: *
20: em_ges_vrt: *
21: em_ges_fun: *
22: em_ges_mam: *
23: em_ges_mus: *
24: em_ges_pro: *
25: em_ges_rtd: *
26: em_ges_pbg: *
27: em_ges_vrl: *
28: gb_ges1: *
29: gb_ges2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	700	37.6	700	12	B1565826	B1565826 RH63652.5
2	684.4	36.7	741	9	AA567307	AA567307 HL01004.5
3	652	35.0	691	12	B1369922	B1369922 RE56261.5
4	640.4	34.4	658	10	BG640863	BG640863 SD12027.5

5	634	34.0	634	12	B1620793	B1620793 RH52016.5
6	630	33.8	630	9	A1402394	A1402394 GH21610.5
7	623	33.4	624	9	A1405571	A1405571 SD25612.3
8	622	33.4	698	12	B1393361	B1393361 SD21925.5
9	611.4	32.8	627	14	CA805481	CA805481 ESC0118b.
10	605	32.5	605	9	A1405070	A1405070 GH24982.5
11	603	32.4	700	12	B1486012	B1486012 RE69377.5
12	598.2	32.1	714	9	AA391125	AA391125 LD10222.5
13	596	32.0	596	12	B1622312	B1622312 RH54158.5
14	590.8	31.7	607	12	B1617864	B1617864 RH48328.5
15	587.4	31.5	592	9	AA141026	AA141026 CK01297.3
16	585.8	31.4	590	9	AA735971	AA735971 GM10221.5
17	581.4	31.2	661	10	BF502519	BF502519 AT18376.5
18	581.4	31.2	673	9	AA950894	AA950894 SD21054.3
19	579	31.1	581	9	A1389866	A1389866 GH21154.3
20	579	31.1	632	12	B1590440	B1590440 RH04062.5
21	573	30.8	676	12	B1579548	B1579548 RE73595.5
22	572.6	30.7	591	12	B1638604	B1638604 SD20938.5
23	572.6	30.7	646	12	B1636537	B1636537 SD18247.5
24	572.6	30.7	647	10	BG637323	BG637323 SD15243.5
25	572.6	30.7	657	12	B1638180	B1638180 SD20405.5
26	570.8	30.6	622	9	AA820761	AA820761 LD24637.3
27	569.2	30.6	588	9	AA697557	AA697557 HL02768.5
28	565	30.3	565	9	A1512604	A1512604 LD44523.5
29	564.4	30.3	632	9	A1109271	A1109271 GH08410.5
30	561	30.1	562	9	A1106957	A1106957 GH06228.5
31	548	29.4	550	12	B1616271	B1616271 RH46174.5
32	547	29.4	549	12	B1166445	B1166445 RE06571.5
33	541.8	29.1	624	12	B1485289	B1485289 RE68591.5
34	541.6	29.0	548	10	BG640701	BG640701 SD11841.5
35	541	29.0	651	9	AA699134	AA699134 HL07773.5
36	539.6	28.9	557	12	B1353505	B1353505 GM23962.5
37	538.6	28.9	636	12	B1358065	B1358065 RE44732.5
38	538	28.9	539	9	A1455243	A1455243 LD22217.3
39	536.4	28.8	549	9	AA390384	AA390384 LD09158.5
40	536.2	28.8	610	9	A1108651	A1108651 GH07860.5
41	535.6	28.7	610	9	A1109037	A1109037 GH07536.5
42	534	28.7	606	9	A1258316	A1258316 LP01471.5
43	533	28.6	591	9	A1388688	A1388688 CH119663.5
44	532	28.6	628	12	B1228490	B1228490 RE26045.5
45	530.8	28.5	546	12	B1615069	B1615069 RH44591.5

ALIGNMENTS

RESULT 1
B1565826
LOCUS
DEFINITION RH63652.5prime RH Drosophila melanogaster normalised Head pfic-1
Drosophila melanogaster cDNA clone RH63652.5 similar to CG8029:
Phan0008029 GO: [enzyme (GO:0003824)] located on: 2R 45B7-45B7.7;
08/23/2001, mRNA sequence.
ACCESSION B1565826
VERSION B1565826.1 GI:15455018
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1. (bases 1 to 700)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farian, D., Frise, B., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mista, S., Mungall, C. J., Nuno, J., Pacled, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celnik, S. and Rubin, G. M.
BDGP/HMI RH Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
 hit genomic AB003834: arm:2R [4092456,4349791]
 estimated-cyto:45A7-45C5: 08/23/2001
 Plate: RH.636 row: E column: 4
 High quality sequence stop: 595.
 Location/Qualifiers

FEATURES

source

1. 700
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="RH63652"
 /sex="male and female"
 /dev_stage="Adult"
 /lab_host="DHS-alpha Tona"
 /clone_id="RH Drosophila melanogaster normalized Head
 p1c-1"
 /note="Organ: head; Vector: p1c1; Site 1: XhoI; Site 2:
 BamHI; Library was kindly generated by Piero Carnicci
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."

BASE COUNT

167 a 182 c 154 g 197 t

ORIGIN

Query Match 37.6%; Score 700; DB 12; Length 700;

Best Local Similarity 100.0%; Pred. No. 3.7e-134; Indels 0; Gaps 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1039 TCCGCGCCGCTCAACACATGTACACACCTGAGCTTCAAGTCCCTGAGCTGAGAGCT 1098
 1 TCCGCGCCGCTCAACACATGTACACACCTGAGCTTCAAGTCCCTGAGCTGAGAGCT 60
 1099 CCCTTCATGTAGTACAGAGAGACTTCCCTTCGCGCACTCTGAGACTGTGTGGCC 1158
 61 CCCTTCATGTAGTACAGAGAGACTTCCCTTCGCGCACTCTGAGACTGTGTGGCC 120
 1159 TTGCTGACGCCCGGATCTGATGAGGAGCTGTTGTGGTCCGCTTCACTGCTGATCATG 1218
 121 TTGCTGACGCCCGGATCTGATGAGGAGCTGTTGTGGTCCGCTTCACTGCTGATCATG 180
 1219 TTGCTGACGCCCGGATCTGATGAGGAGCTGTTGTGGTCCGCTTCACTGCTGATCATG 1278
 181 TTGCTGACGCCCGGATCTGATGAGGAGCTGTTGTGGTCCGCTTCACTGCTGATCATG 240
 1279 GGCAGAGCTATCAACATCAATGCGCGCGAGTAAACCGCTTCACTGCTGATCATG 1338
 241 GGCAGAGCTATCAACATCAATGCGCGCGAGTAAACCGCTTCACTGCTGATCATG 300
 1339 CTTATTTGATTTGATTCATGAGGAGCTCAATATTCCTTATTCCTTCACTGATCATG 1398
 301 CTTATTTGATTTGATTCATGAGGAGCTCAATATTCCTTATTCCTTCACTGATCATG 360
 1399 ATGTATATGTTAAAGATCAATATCTTAATACCGTACACGATTAACATTAACATTAAC 1458
 361 ATGTATATGTTAAAGATCAATATCTTAATACCGTACACGATTAACATTAACATTAAC 420
 1459 GGGAAATTAAGTCTTCACAATTTGATGATTCGCTTAATTTTGGCGCGAGAGATCAATT 1518
 421 GGGAAATTAAGTCTTCACAATTTGATGATTCGCTTAATTTTGGCGCGAGAGATCAATT 480
 1519 ACCTTGACGATCAAGGTTGCTGATCTTGGCGCGAGAAATTTGGCGAGAGATTTGCTT 1578
 481 ACCTTGACGATCAAGGTTGCTGATCTTGGCGCGAGAAATTTGGCGAGAGATTTGCTT 540
 1579 GCAACGAGCGAGATCAAGTATTCATTTAGATAGAGCCCTTAATGTCGTTATCGAAG 1638
 541 GCAACGAGCGAGATCAAGTATTCATTTAGATAGAGCCCTTAATGTCGTTATCGAAG 600
 1639 AACTTTAAAGAGCGACATGTGTGTGTGTTTAAAGAGCGCGCTGTGTGTGTGTGCA 1698
 601 AACTTTAAAGAGCGACATGTGTGTGTGTTTAAAGAGCGCGCTGTGTGTGTGTGCA 660

QY 1699 CATGGGAGTAATCTAATCCGAGTATATCTAAGTGT 1738
 DB 661 CATGGGAGTAATCTAATCCGAGTATATCTAAGTGT 700

RESULT 2
 AA567307 741 bp mRNA linear EST 19-APR-2001
 LOCUS HL01004.5prime HL Drosophila melanogaster head Bluescript
 DEFINITION Drosophila melanogaster cDNA clone HL01004 5prime, mRNA sequence.

AA567307
 AA567307.1 GI:2639888
 EST.
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 741)
 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
 Lewis, S. and Rubin, G. M.
 BDGP/HMI Drosophila EST Project
 Unpublished
 On Dec 18, 1997 this sequence version replaced gi:2339819.
 Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
 Plate: 10 row: A column: 4
 High quality sequence stop: 667.
 Location/Qualifiers

FEATURES

source

1. 741
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="BDGP_EST:BDClnd02023"
 /db_xref="taxon:7227"
 /clone="HL01004"
 /sex="male and female"
 /dev_stage="adult"
 /lab_host="SOLR"
 /clone_id="HL Drosophila melanogaster head Bluescript"
 /note="Organ: head--brain & sensory organ; Vector:
 Bluescript SK; Site 1: EcoRI; Site 2: XhoI; Constructed
 using Stratagene ZAP-cDNA Synthesis kit. Oligo dt-primed
 and directionally cloned at EcoRI and XhoI in Bluescript
 SK(+/-)"

BASE COUNT 151 a 238 c 183 g 167 t 2 others

ORIGIN

Query Match 36.7%; Score 684.4; DB 9; Length 741;

Best Local Similarity 98.3%; Pred. No. 6.2e-131; Indels 5; Gaps 4;
 Matches 733; Conservative 0; Mismatches 8; Indels 5; Gaps 4;

527 GAAAGTGACGACGCGGCAAGTGTGCGTATGAGCGGCTGTGTTGATGACCTTGAGG 586
 1 GAAAGTGACGACGCGGCAAGTGTGCGTATGAGCGGCTGTGTTGATGACCTTGAGG 58
 587 ACCCGCGGAGAGCCGAGGAGGCGATCTGAGTCAATGACGCGGCATATGCTTATCA 646
 DB 59 ACCCGCGGAGAGCCGAGGAGGCGATCTGAGTCAATGACGCGGCATATGCTTATCA 117
 647 GCAAGAGTTGAGTGCAAGTGTGCTTATGTAATCTGAGCGGCGGCTTCAAGGCTTCG 706
 118 GCAAGAGTTGAGTGCAAGTGTGCTTATGTAATCTGAGCGGCGGCTTCAAGGCTTCG 177
 707 TGGTGAGGCTGTATCCCGCGTATGACACGCGCCCAACCGGTGATCATGTGAAGT 766
 DB 178 TGGTGAGGCTGTATCCCGCGTATGACACGCGCCCAACCGGTGATCATGTGAAGT 237
 767 CGAACATCATTTTCAATCTTATCACTGCTGTCTTAAACGAGCAACCCCATCAACG 826

Db 238 CGAACCAATGATTTAGATCTTCTACAGTCCCTGCTTACAAACGGAACCCCATCACCG 297
Qy 827 TCACGACCTGAGCTACCAACTAGCTTACCAAGCTATCCGTTTATGACACAT 886
Db 298 TCACGACCTGAGCTACCAACTAGCTTACCAAGCTATCCGTTTATGACACAT 357
Qy 887 CTGTGCGCAGCAAGCCCAATCAGCTTGTGTTTATATATGCGGATACCTTAGTCTGA 946
Db 358 CTGTGCGCAGCAAGCCCAATCAGCTTGTGTTTATATATGCGGATACCTTAGTCTGA 417
Qy 947 GCATTTTGTCTACGACAAACAATTCCTGCTTACGCGGCGGTGATGCCCCCAACGAT 1006
Db 418 GCATTTTGTCTACGACAAACAATTCCTGCTTACGCGGCGGTGATGCCCCCAACGAT 477
Qy 1007 TCTCTTACTGCTGCGGCAACCTCAGCCCTGAGTCCGCGGCGGTGATGATGATACAA 1066
Db 478 TCTCTTACTGCTGCGGCAACCTCAGCCCTGAGTCCGCGGCGGTGATGATGATACAA 537
Qy 1067 CCTGAGCTTCAAGTCCCTGAGCTGAGGCTCCCTTCAATGATGATGATGATGATGAT 1126
Db 538 CCTGAGCTTCAAGTCCCTGAGCTGAGGCTCCCTTCAATGATGATGATGATGATGAT 596
Qy 1127 TTCCTTGGGCACTCTCGGAGCTGTGTGAGTGTGTGAGTGTGTGAGTGTGTGAGTGT 1186
Db 597 TTCCTTGGGCACTCTCGGAGCTGTGTGAGTGTGTGAGTGTGTGAGTGTGTGAGTGT 656
Qy 1187 TGTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246
Db 657 TGTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
Qy 1247 TCACACGATGAGACCGCTTCGACGAT 1272
Db 716 TCACACGATGAGACCGCTTCGACGAT 741

RESULT 3
LOCUS B1369922 691 bp mRNA linear EST 01-AUG-2001
DEFINITION R556261.5prime RE Drosophila melanogaster normalized Embryo pFLC-1
Drosophila melanogaster cDNA clone R556261.5 similar to CG8029:
FBa0008029 GO: [enzyme (GO:0003824)] located on: 2R 45B7-45B7.1;
05/15/2001, mRNA sequence.

ACCESSION B1369922
VERSION B1369922.1 GI:15065950

KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster
Neoptera; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Ephedroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 691)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson,
J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Fries, E., George,
R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mistr, S.,
Mungall, C. J., Nuno, J., Pacle, J., Paragas, V., Park, S.,
Phanphanavong, S., Wan, K., Yu, C., Lewis, S. B., Celniker, S. and Rubin,
G. M.

TITLE BDGP/HMT RE Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AB003834: arm:2R [4092456, 4349779]
estimated-cyto:45A7-45C5: 05/15/2001
Plate: RE 562 row: F column: 1
High quality sequence stop: 568.

FEATURES
source
1. 691
/organism="Drosophila melanogaster"
/mol_type="mRNA"

/db_xref="taxon:7227"
/clone="R556261"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tora"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFLC-1"
/note="Organ: embryo; Vector: pFLC1; Site 1: XhoI; Site 2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 140 a 189 c 194 g 168 t
ORIGIN

Query Match 35.0%; Score 652; DB 12; Length 691;
Best Local Similarity 99.8%; Pred No. 36-124;
Matches 663; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CGACGTGAGCAAGGCAAGCAAAATTTTCAACACATTTTGTGTTAGTTTCAATT 60
Db 29 CGACGTGAGCAAGGCAAGCAAAATTTT-ACACACATTTTGTGTTAGTTTCAATT 87
Qy 61 CACTTTGCGTGTGTTTATGCTCTTGTGCTTCTCTTGTGCTTATGATTT 120
Db 88 CACTTTGCGTGTGTTTATGCTCTTGTGCTTCTCTTGTGCTTATGATTT 147
Qy 121 TGTGTTTCTTGTGTTTATGCTCTTGTGCTTCTCTTGTGCTTATGATTT 180
Db 148 TGTGTTTCTTGTGTTTATGCTCTTGTGCTTCTCTTGTGCTTATGATTT 207
Qy 181 TCGAATATGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 208 TCGAATATGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 267
Qy 241 TTTTGTGAGGAGCAAGCAAGTGTGCAAACTCTCTGAGACGCTGCTCAAGTGAG 300
Db 268 TTTTGTGAGGAGCAAGCAAGTGTGCAAACTCTCTGAGACGCTGCTCAAGTGAG 327
Qy 301 TTTTGTGAGGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 328 TTTTGTGAGGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
Qy 361 GGCCTGAGCAAGGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 388 GGCCTGAGCAAGGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
Qy 421 CTGCAAGGAGTGAAGCCCAAGCTTACTACACAGCTGAGAGACCTCTCGAGGCACTG 480
Db 448 CTGCAAGGAGTGAAGCCCAAGCTTACTACACAGCTGAGAGACCTCTCGAGGCACTG 507
Qy 481 CGCTCAGTGGCCGCAAGGCAAGCACTTCAATGAGTGAAGTGAAGTGAAGTGAAGTGA 540
Db 508 CGCTCAGTGGCCGCAAGGCAAGCACTTCAATGAGTGAAGTGAAGTGAAGTGAAGTGA 567
Qy 541 CCGGCAAGTGTGCGTGAAGTGAAGGCTGTTTGTGATCTTGAAGACCGCCCAAGAGC 600
Db 568 CCGGCAAGTGTGCGTGAAGTGAAGGCTGTTTGTGATCTTGAAGACCGCCCAAGAGC 627
Qy 601 CGGAGGCGAGTCTGAGTCAATGACGCGCATAGCTTATACGAAGAGTTTGAAG 660
Db 628 CGGAGGCGAGTCTGAGTCAATGACGCGCATAGCTTATACGAAGAGTTTGAAG 687
Qy 661 TCGA 664
Db 688 TCGA 691

RESULT 4
LOCUS BG640863 658 bp mRNA linear EST 23-APR-2001
DEFINITION SD12027.5prime SD Drosophila melanogaster Schneider I2 cell culture
PORT2 Drosophila melanogaster cDNA clone SD12027.5 similar to
CG8029: FBa0008029 located on: 2R 45B7-45B7.1; 04/13/2001, mRNA

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BG640863	BG640863			Drosophila melanogaster (fruit fly)					
BG640863.1	GI:13772789			Drosophila melanogaster (fruit fly)					
EST.				Drosophila melanogaster					
				Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridae; Drosophilidae; Drosophila.					
				1 (bases 1 to 658)					
				Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S., and Rubin, G.M.					
				BDBP/HMI Drosophila EST Project					
				Unpublished					
				Contact: Stapleton, M.					
				BDBP					
				Lawrence Berkeley National Lab					
				Ome Cyclotron Rd, Berkeley, CA 94720, USA					
				Fax: 510 486 6798					
				Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu					
				hit genomic AB003834 : arm:2R [4092456,4349779]					
				estimated-cyto:45A7-45C5 : 04/13/2001					
				Plate: SD 120 row: C column: 3					
				High quality sequence stop: 569.					
				Location/Qualifiers					
				1..658					
				/organism="Drosophila melanogaster"					
				/mol_type="mRNA"					
				/db_xref="taxon:7227"					
				/clone="SD12027"					
				/lab_host="DH5-alpha"					
				/clone_lib="SD drosophila melanogaster Schneider L2 cell culture POT2"					
				/note="Vector: POT2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into POT2."					
				Plasmid cDNA library." 8 others					
				BASE COUNT					
				ORIGIN					
				132 a 179 c 186 g 153 t					
				Query Match					
				Best Local Similarity 97.9%; Pred. No. 7.3e-122;					
				Matches 644; Conservative 0; Mismatches 14; Indels 0; Gaps 0,					
				34.4%; Score 640.4; DB 10; Length 658;					
				1 CGAGCTCGAGCAAGCAAAGCAAATTTTACCACCATTTTTTGTTGTAGTTTCAGTT					
				61 CACTTTGGGTTGGTTTAAAGGCGCTCTGCGTTCTCCCTGGTTCTTAATTATCATTT					
				61 CACTTTGGGTTGGTTTAAAGGCGCTCTGCGTTCTCCCTGGTTCTTAATTATCATTT					
				121 TTGCTTCCTTTGTTTTCATTAAACCCCACCGAAGTAGCGAATCCAGCGCGATGTTG					
				121 TTGCTTCCTTTGTTTTCATTAAACCCCACCGAAGTAGCGAATCCAGCGCGATGTTG					
				181 TGGAAATGCGATGATGCCGTGTCGTCATGCGGAGCGCGTGCAGGAAGAAGCGCGGTC					
				241 TTTTGTGGGGAGCAACAGTGTGGCGAAACCTCTCGTAGAACGCGTGTCCCAATGGAG					
				241 TTTTGTGGGGAGCAACAGTGTGGCGAAACCTCTCGTAGAACGCGTGTCCCAATGGAG					
				301 TTGCGCAGACAGTGGCTGATGTTGCTGGAAGATCATATGTCGTGGCTTTGAGGAATAAT					
				301 TTGCGCAGACAGTGGCTGATGTTGCTGGAAGATCATATGTCGTGGCTTTGAGGAATAAT					
				361 GGCCGAGAGAGAAAGATCTCTGTCGCTCAATCCCAAGCGCAATCTCTGTCAGCCGAG					
				361 GGCCGAGAGAGAAAGATCTCTGTCGCTCAATCCCAAGCGCAATCTCTGTCAGCCGAG					
				421 GTGACGGAGTAGGCCCAAGACTTACAACACAGCTGAGAAACCTCTGGAGGCACTG					

Db	CTGACGGAGTGTAGAGCCCAAGACTACTATACACGACGGTGGAGAAACCCCTCGAGAGGCACTG	480
Qy	481 CGCTCACTGGCCGCCCAAGCGCGGAGACAACTCCATCGATGCGCACTGGAAAGCTGACCAAG	540
Db	481 CGCTCGATGTGNCGCAACGCGGAGACAACTCCATCGATGCGCAAGGAAAGCTGACCAAG	540
Qy	541 CCGGCGCAAGTGTGCGCGAGGTACGGGCGCTGTGTGTACCTTCGAGAGACGCGCGGACGAGC	600
Db	541 CCGGNCAAAGTGTGCGCGTGTAGTACGAGNCCTGTGTGTACCTTCGAGAGACGCGCGNCGAGAG	600
Qy	601 CCGGAGGCGCACTGTGATGCATATGACGCGCGCATAGCTCTATCAGACAGCAAGTTTCG	658
Db	601 CCGGAGGCGCAAGNCTGAGTGCATATGACGCGCGCATAGCTCTATCAGACAGCAAGCTTCG	658
RESULT 5		
LOCUS	B1620793	
DEFINITION	B1620793 634 bp mRNA linear EST 07-SEP-2001	
ACCESSION	RH52016 5prime RH Drosophila melanogaster normalized Head p1c-1	
VERSION	Drosophila melanogaster cDNA clone RH52016 5 similar to CG8029:	
KEYWORDS	Fan0008029 GO:enzyme (GO:0003824) Located on: 2R 45B1-45B77.1	
ORGANISM	08/22/2001, mRNA sequence.	
REFERENCE	B1620793 GI:15516318	
AUTHORS	EST.	
DESCRIPTION	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
AUTHORS	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
DESCRIPTION	Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 1 to 634)	
AUTHORS	Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson	
DESCRIPTION	, J., Champagne, M., Chavez, C., Dorsett, V., Fattah, D., Frise, E., George	
AUTHORS	, R., Gonzalez, M., Guarini, H., Harris, N., Li, P., Liao, G., Mista, S.,	
DESCRIPTION	Mungall, C.J., Nunoo, J., Pacleb, D., Pargae, V., Park, S.,	
AUTHORS	Poulsen, S., Wan, K., Yu, C., Lewis, S.E., Celiker, S. and Rubin	
DESCRIPTION	, G.M.	
REFERENCE	BDGP/HMT RH Drosophila EST Project	
AUTHORS	Unpublished	
DESCRIPTION	Contact: Stapleton, M.	
REFERENCE	BDGP	
DESCRIPTION	Lawrence Berkeley National Lab	
AUTHORS	One Cyclotron Rd, Berkeley, CA 94720, USA	
DESCRIPTION	Fax: 510 486 6798	
AUTHORS	Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu	
DESCRIPTION	hit genomic AB003834: arm:2R [4092456..4349779]	
AUTHORS	estimated-cyto:45A7-45C5: 08/22/2001	
DESCRIPTION	Plate: RH:520 row: B column: 4	
AUTHORS	High quality sequence stop: 632.	
DESCRIPTION	Location/Qualifiers	
REFERENCE	1..634	
DESCRIPTION	/organism="Drosophila melanogaster"	
AUTHORS	/mol_type="mRNA"	
DESCRIPTION	/db_xref="taxon:7227"	
AUTHORS	/clone="RH52016"	
DESCRIPTION	/sex="male and female"	
AUTHORS	/dev_stage="Adult"	
DESCRIPTION	/lab_host="RH5-alpha Tona"	
AUTHORS	/clone_lib="RH Drosophila melanogaster normalized Head	
DESCRIPTION	p1c-1"	
AUTHORS	/note="Organ: head, Vector: p1c1, Site 1: XhoI; Site 2:	
DESCRIPTION	BamHI; Library was kindly generated by Piero Carninci at	
AUTHORS	the RIKEN. The library was normalized and excised using	
DESCRIPTION	Cre recombinase. Plasmid cDNA library."	
BASE COUNT	151 a 171 c 134 g 178 t	
ORIGIN		
Query Match	34.0%; Score 634; DB 12; Length 634;	
Best Local Similarity	100.0%; Pred. No. 1.5e-120;	
Matches	634; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
1039	TCGCGGCGCGTCAACAATGTACACACCCCTGAGCTTCAAGTCCCTGACAGCTCGAGGCT	1098

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT										
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	601
Db	1	1009	61	1159	121	1219	181	1279	241	1339	301	1399	361	1459	421	1519	481	1579	541	1639	6

	/db xref="taxon:7227"	
	/clone="GH21610"	
	/sex="male and female"	
	/dev_stage="adult"	
	/lab_host="DHS - alpha"	
	/clone_id="GH Drosophila melanogaster head pot2"	
	/note="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2:	
	XhoI; Sized fractionated cDNAs were directly ligated into	
	pot2. Plasmid cDNA library."	
BASE COUNT	131 a 208 c 144 g 147 t	
ORIGIN		
Query Match	33.8%; Score 630; DB 9; Length 630;	
Best Local Similarity	100.0%; Prod. No. 1e-119;	
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Dy	CTTATCTGTACTGCGCCGCCCTTCACAGCGCTCCTGTGTGCACGGTGTAACCCGCGTG	730
Db	1 CTTATCTGTACTGCGCCGCCCTTCACAGCGCTCCTGTGTGCACGGTGTAACCCGCGTG	60
Dy	731 ACACCGCGGCACAACCGGTGGCATCATGTGGAAAGTGAACCAATCAGTTTCATCTTCT	790
Db	61 ACACCGCGGCACAACCGGTGGCATCATGTGGAAAGTGAACCAATCAGTTTCATCTTCT	120
Dy	791 ACAAAGCGCGCTTCAACAAGCGCAACCCCAATCACACGTACACGACCTCAAGTCCCACT	850
Db	121 ACAAAGCGCGCTTCAACAAGCGCAACCCCAATCACACGTACACGACCTCAAGTCCCACT	180
Dy	851 CTAGCTCTACCAAGCTATCCGTTTATTATGACATCTGTGCGCGACAGCAAGCAATCACT	910
Db	181 CTAGCTCTACCAAGCTATCCGTTTATTATGACATCTGTGCGCGACAGCAAGCAATCACT	240
Dy	911 TTGACGTTGTTTATAATGCGGAGATCTTTAAGTCTGAGCAATTTGTTCTTAGACAACA	970
Db	241 TTGACGTTGTTTATAATGCGGAGATCTTTAAGTCTGAGCAATTTGTTCTTAGACAACA	300
Dy	971 ACTTCCGCTCAGGCGGGGTGAATCCCCAACACAGTCTCTTACTCCGCGCGCAACTCA	1030
Db	301 ACTTCCGCTCAGGCGGGGTGAATCCCCAACACAGTCTCTTACTCCGCGCGCAACTCA	360
Dy	1031 CCCTCGAGTCCGCGCGCGCTCAACAACATGTACACACCCTGAGCTTCAAGTCCCTGAGC	1090
Db	361 CCCTCGAGTCCGCGCGCGCTCAACAACATGTACACACCCTGAGCTTCAAGTCCCTGAGC	420
Dy	1091 TGCAGGCTCCCTTCGATGATACGTACAGAAGAGACTTTCCTTCGCGGACTCTCGGACT	1150
Db	421 TGCAGGCTCCCTTCGATGATACGTACAGAAGAGACTTTCCTTCGCGGACTCTCGGACT	480
Dy	1151 GTGTGGGCTTGTGTACACCGCGGTATCCTGATGGGACTGTTTGTGTGTGCGCTTACTGTG	1210
Db	481 GTGTGGGCTTGTGTACACCGCGGTATCCTGATGGGACTGTTTGTGTGTGCGCTTACTGTG	540
Dy	1211 TCATCATGTTGTGTGGGCGTCTGCGATGATGACATCAACAGATGGAACCGTTCGACG	1270
Db	541 TCATCATGTTGTGTGGGCGTCTGCGATGATGACATCAACAGATGGAACCGTTCGACG	600
Dy	1271 ATCCCAAGGGCAGACTATCACCATCAATG 1300	
Db	601 ATCCCAAGGGCAGACTATCACCATCAATG 630	
RESULT 7		
A1405571/c	624 bp mRNA linear EST 19-APR-2001	
LOCUS	GH25612.3prIME GH Drosophila melanogaster head pot2 Drosophila	
DEFINITION	melanogaster cDNA clone GH25612 3prime, mRNA sequence.	
ACCESSION	A1405571	
VERSION	A1405571.1 GI:4248658	
KEYWORDS	EST.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Diptera; Eudiptera; Brachycera; Muscomorpha; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	

REFERENCE
1 (bases 1 to 624)
AUTHORS
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G. M.
TITLE
BDGF/HNMI Drosophila EST Project
JOURNAL
Unpublished
COMMENT
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone
probably contains an inverted insert. The resulting Poly-T sequence
has been removed.
Plate: 256 row: A column: 12
High quality sequence stop: 525.
Location/Qualifiers
1. 624
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH25612"
/sex="male and female"
/dev_stage="adult"
/lab_host="DHS - alpha"
/clone_lib="GH Drosophila melanogaster head pot2"
/note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pot2. Plasmid cDNA library."
BASE COUNT
194 a 120 c 145 g 165 t
ORIGIN
Query Match 33.4%; Score 623; DB 9; Length 624;
Best Local Similarity 100.0%; Pred. No. 2,8e-118;
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1241 TGGACATCAACAGATGACCGCTTGCAGCTCCAGGCAAGACTATCACCATCANTG 1300
624 TGGACATCAACAGATGACCGCTTGCAGCTCCAGGCAAGACTATCACCATCANTG 565
1301 CCGCGCGCGATAAACCGCTACATCATCGCTTGCACACTTATTTGATTGCTCAGTGC 1360
564 CCGCGCGCGATAAACCGCTACATCATCGCTTGCACACTTATTTGATTGCTCAGTGC 505
1361 ACTCAATATTCATCTTATTCCTTCAAGTCTATCCACATGATATGTTAAAGATCAAT 1420
504 ACTCAATATTCATCTTATTCCTTCAAGTCTATCCACATGATATGTTAAAGATCAAT 445
1421 ACTTAATACCGTACAGCATTAACATTAACATTAACCGGGAATTAGTCTCCATTTT 1480
444 ACTTAATACCGTACAGCATTAACATTAACATTAACCGGGAATTAGTCTCCATTTT 385
1481 GTATCGATCCGTAATTTTGCAGCAAGATCAGTTTACGTTGAGATCACAAGCTTGC 1540
384 GTATCGATCCGTAATTTTGCAGCAAGATCAGTTTACGTTGAGATCACAAGCTTGC 325
1541 TGATTTCTGTGCGCGCAAAATTTGCTTGTGCTTGCACGAGCGAGATCAGTAT 1600
324 TGATTTCTGTGCGCGCAAAATTTGCTTGTGCTTGCACGAGCGAGATCAGTAT 265
1601 TTCAATTTAGATACCCCTTAATGCTTATTCGAAGAACTTTAAAGACGACATGTG 1660
264 TTCAATTTAGATACCCCTTAATGCTTATTCGAAGAACTTTAAAGACGACATGTG 205
1661 TGTGTGTGTTAAGAGCGCGCTGTGTGTGTGTCAGATGCGGCAAGTAATCAATCCG 1720
204 TGTGTGTGTTAAGAGCGCGCTGTGTGTGTGTCAGATGCGGCAAGTAATCAATCCG 145
1721 AAGTAATCTAAGTGTATTAAGTTCTTGTGTTGCGCAAAATATGCGAGACTAATTTATG 1780
144 AAGTAATCTAAGTGTATTAAGTTCTTGTGTTGCGCAAAATATGCGAGACTAATTTATG 85

RESULT 8
BI639361 698 bp mRNA linear EST 10-SEP-2001
LOCUS
SD21925.5prime SD Drosophila melanogaster Schneider L2 cell culture
DEFINITION
SD21925.5prime SD Drosophila melanogaster cDNA clone SD21925 5 similar to
CG8029; FBan008029 GO: [enzyme (GO:0003824)] located on: 2R
45B7-45B7; 05/19/2001, mRNA sequence.
ACCESSION
BI639361 GI:15541571
VERSION
BI639361.1
KEYWORDS
EST.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 698)
AUTHORS
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G. M.
TITLE
BDGF/HNMI Drosophila EST Project
JOURNAL
Unpublished
COMMENT
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hlt genomic AB003834; arm: 2R [4092456, 4349779]
estimated-cyto: 45A7-45C5; 05/19/2001
Plate: SD.219 row: C column: 1
High quality sequence stop: 596.
Location/Qualifiers
1. 698
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="SD21925"
/lab_host="DHS-alpha"
/clone_lib="SD Drosophila melanogaster Schneider L2 cell
culture pot2"
/note="Vector: pot2; Site_1: EcoRI; Site_2: XhoI; Sized
fractionated cDNAs were directly ligated into pot2.
Plasmid cDNA library."
BASE COUNT
145 a 207 c 203 g 143 t
ORIGIN
Query Match 33.4%; Score 622; DB 12; Length 698;
Best Local Similarity 99.2%; Pred. No. 4.6e-118;
Matches 625; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
157 GTAAGCGAATCCAGCGGATGTTGTGAATTCGTTGCTGCTGCTATTTGGGCT 216
69 GTAAGCGAATCCAGCGGATGTTGTGAATTCGTTGCTGCTGCTATTTGGGCT 128
217 GCCGTGGCGGAGCAACCGCCGCTTTTGTGGGAGCCCAAGTGGGGAACCTCC 276
129 GCCGTGGCGGAGCAACCGCCGCTTTTGTGGGAGCCCAAGTGGGGAACCTCC 188
277 CTGAAGCGGTGTCCTCAAGTGAAGTTCGAGACAGTGGCTGATTCGTAAGATCAC 336
189 CTGAAGCGGTGTCCTCAAGTGAAGTTCGAGACAGTGGCTGATTCGTAAGATCAC 248
337 ATGTCGTGACCTTCGAGAAATGCGCTGAGAGCAAGACTTCTGTGCTCAATCC 396

Db 249 ATGCTGTGGCTTTCAGAGAAAATGGCTGAGACAGAGACTTCTGTGCTCAACTCC 308
QY 397 CAGGCCAGTCTGCTACGCCAGCTGACAGGAGTAGAGCCCAAGACTTACTACACAG 456
Db 309 CAGGCCAGTCTGCTACGCCAGCTGACAGGAGTAGAGCCCAAGACTTACTACACAG 368
QY 457 GTGAGAAACCTCTGAGAGGACCTGCGTCAGTGGCCGCCAAGCGAGACAACTCCATC 516
Db 369 GTGAGAAACCTCTGAGAGGACCTGCGTCAGTGGCCGCCAAGCGAGACAACTCCATC 428
QY 517 GATGCAAGTGAAGAGTGAACAGCGCGGCAATGTGCGCTAGTACGCGCTTTTGG 576
Db 429 GATGCAAGTGAAGAGTGAACAGCGCGGCAATGTGCGCTAGTACGCGCTTTTGG 488
QY 577 ACCTTGAGAGACGCCCGGAGAGCCGGAGAGCCAGTCTGAGTCAATGACGCCGCAT 636
Db 489 ACCTTGAGAGACGCCCGGAGAGCCGGAGAGCCAGTCTGAGTCAATGACGCCGCAT 548
QY 637 GCTGCTATCAGCAGAGAGTTCAGAGTGCAGAGTGGCTTATCTGACCTGGCCCTCC 696
Db 549 GCTGCTATCAGCAGAGAGTTCAGAGTGCAGAGTGGCTTATCTGACCTGGCCCTCC 608
QY 697 AGCGCTCTCTGTGTGACAGCTGCTACCCCGCTGACACCGCGCCCAACCGGTGGCAT 756
Db 609 ACCGCTCTCTGTGTGACAGCTGCTACCCCGCTGACACCGCGCCCAACCGGTGGCAT 668
QY 757 ATGTGAGAGTGCACCAATCAGTTTCAGATC 786
Db 669 ATGTGAGAGTGCACCAATCAGTTTCAGATC 698

RESULT 9
CA805481/c 627 bp mRNA linear EST 09-DEC-2002
LOCUS ES60118B.B21.P09.3prime ES601 Drosophila melanogaster cDNA 3'

DEFINITION
CA805481
CA805481
CA805481.1 GI:26254430
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 627)
Gorski,S.M., Chittaranjan,S., Pleasance,E.D., Freeman,J.D.,
Anderson,C.L., Varhol,R.J., Coughlin,S.M., Zuyderduyn,S.D., Jones
S.J.M. and Marra,M.A.
A SAGE Approach to Discovery of Genes Involved in Autophagic Cell
Death
Curr. Biol. 13 (4), 358-363 (2003)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Genome Sciences Centre
British Columbia Cancer Agency
600 West 10th Ave., Room 3427, Vancouver, BC, Canada, V5Z 4E6
Tel: 604 877 6081
Fax: 604 877 6085
Email: sgorski@bcgsc.bc.ca
Seq primer: -21M13F.
Location/Qualifiers
1. 627
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="Oregon R"
/db_xref="taxon:7227"
/sex="male and female"
/tissue_type="salivary gland"
/dev_stage="16, 18, 20, 22, and 24 hrs after puparium
formation at 18°C"
/lab_host="DH108"
/clone_lib="ES601"

FEATURES
source

/note="Vector: pSPOR1; Site 1: NotI; Site 2: SalI; mRNA
isolated from 500 pairs of hand-dissected salivary glands
from mixed stage animals was used to construct an
oligo-dT-primed directional cDNA library"
BASE COUNT 198 a 122 c 142 g 165 t
ORIGIN

Query Match 32.8%; Score 611.4; DB 14; Length 627;
Best Local Similarity 99.0%; Pred. No. 7.1e-116;
Matches 615; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1237 ATGATGACATCAACAGATGAGACCGCTTGCAGATCCCAAGGCAAGCTATCACATC 1296
Db 627 ATGATGACATCAACAGATGAGACCGCTTGCAGATCCCAAGGCAAGCTATCACATC 568
QY 1297 AATGCGCGCGGAGTAAACCGCTCAATCAATGATGATCAACCTTATTTGATTTCAAT 1356
Db 567 AATGCGCGCGGAGTAAACCGCTCAATCAATGATGATCAACCTTATTTGATTTCAAT 508
QY 1357 GCGCACTCAATATTCATTCCTTATTCCTTCACTGATCAGATGATATTTAAAGAAAT 1416
Db 507 GCGCACTCAATATTCATTCCTTATTCCTTCACTGATCAGATGATATTTAAAGAAAT 448
QY 1417 CAATACCTTAATACCGTACAGCATTAACATTAACATTAACCGGAAATTAATCTTCA 1476
Db 447 CAATACCTTAATACCGTACAGCATTAACATTAACATTAACCGGAAATTAATCTTCA 388
QY 1477 ATTTGATGATGATCCGTAATTTTCCCGCGCAAGATCAATTAAGTTCAGATCAGAG 1536
Db 387 ATTTGATGATGATCCGTAATTTTCCCGCGCAAGATCAATTAAGTTCAGATCAGAG 328
QY 1537 TTGCTGATTCCTGTGCGCGCAAAATTCCTGAGTTCCTTGCACCCGAGCATCA 1596
Db 327 TTGCTGATTCCTGTGCGCGCAAAATTCCTGAGTTCCTTGCACCCGAGCATCA 268
QY 1597 GATTTCAATTTAAGATAGCCCTTAATGCTGTTATGCAAGAACTTTAAAGACGACA 1656
Db 267 GATTTCAATTTAAGATAGCCCTTAATGCTGTTATGCAAGAACTTTAAAGACGACA 208
QY 1657 TGTGTGTGTGTTTAAAGCGCGCGGTGTGTGTGTGCAATGCGGAGATTTCTTA 1716
Db 207 TGTGTGTGTGTTTAAAGCGCGCGGTGTGTGTGTGCAATGCGGAGATTTCTTA 148
QY 1717 TCCGAAGTATCTAAGTTCATTTATAGTTTCTGTTGGCGCAATATGCGAGCATTA 1776
Db 147 TCCGAAGTATCTAAGTTCATTTATAGTTTCTGTTGGCGCAATATGCGAGCATTA 88
QY 1777 TATGCTCTGATGATAGTTCATTTGTCATTTGCAACCTCTGTTGCGAGGCCCAAG 1836
Db 87 TATGCTCTGATGATAGTTCATTTGTCATTTGTCACACCTCTGTTGCGAGGCCCAAG 28
QY 1837 AAAATTAAGTTTTCATTAAT 1857
Db 27 AAAATTAAGTTTTCATTAAT 7

RESULT 10
A1405070 605 bp mRNA linear EST 19-APR-2001
LOCUS GH24982.5prime GH Drosophila melanogaster head P072 Drosophila
DEFINITION
A1405070
A1405070.1 GI:4248157
EST.
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 605)
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project

JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Plate: 249 row: G column: 10
High quality sequence stop: 487.
Location/Qualifiers

FEATURES
source
1..605
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH24982"
/sex="male and female"
/dev_stage="adult"
/lab_host="DHS - alpha"
/note="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."

BASE COUNT 120 a 170 c 170 g 145 t

ORIGIN
Query Match 32.4%; Score 605; DB 9; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.5e-114;
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGCTGAGCAAGCAAGCAAAATTTTCACCACTTTTCGTTGATGTTCACTT 60
DB 1 CGAGCTGAGCAAGCAAGCAAAATTTTCACCACTTTTCGTTGATGTTCACTT 60
QY 61 CACTTTTCGTTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 CACTTTTCGTTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 TTGCTTCTCTTTGTTTCCATTAAACCCCAACCAAGTAAGCAATCCAGCGCATTTG 180
DB 121 TTGCTTCTCTTTGTTTCCATTAAACCCCAACCAAGTAAGCAATCCAGCGCATTTG 180
QY 181 TGGAAATCGTGTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 TGGAAATCGTGTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TTTTGTGGGAGCAAGCAAGTGGCGAAACCCCTGTAAGCGGTGCCAAGTGGAG 300
DB 241 TTTTGTGGGAGCAAGCAAGTGGCGAAACCCCTGTAAGCGGTGCCAAGTGGAG 300
QY 301 TTTGCGAGCAGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 TTTGCGAGCAGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 GGCCTGAGCAAGCAAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 GGCCTGAGCAAGCAAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 CTGAGAGGAGTGAAGCCCAAGACTTACTACACAGCGTGGAGAACCCCTCGGAGGCACTG 480
DB 421 CTGAGAGGAGTGAAGCCCAAGACTTACTACACAGCGTGGAGAACCCCTCGGAGGCACTG 480
QY 481 CGCTCAGTGGCCGCAAGCGGAGCAACAATCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 CGCTCAGTGGCCGCAAGCGGAGCAACAATCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 CCGGCAAGTGTGCGTAGTACGCGCTGTTTGTGACCTTTCGAGGAGCGCGCGAGAGC 600
DB 541 CCGGCAAGTGTGCGTAGTACGCGCTGTTTGTGACCTTTCGAGGAGCGCGCGAGAGC 600
QY 601 CCGGA 605
DB 601 CCGGA 605

RESULT 11
BI486012
LOCUS
DEFINITION R669377.5p10 RE Drosophila melanogaster normalized Embryo pF1c-1
Drosophila melanogaster cDNA clone R669377.5 similar to CG8029;
F80008029 GO: [enzyme (GO:0003824)] located on: 2R 45B7-45B7.1;
05/16/2001, mRNA sequence.

ACCESSION BI486012
VERSION BI486012.1
KEYWORDS GI:15325793
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 700)

REFERENCE
AUTHORS Stapleton, M., Brockstein, P., Hong, L., Tyler, D., Berman, B., Carlson, J., Chazale, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mista, S., Mungall, C. J., Nunoo, J., Pacle, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Ceiniker, S. and Rubin, G. M.
TITLE BDGP/HMI RE Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
hit genomic AB003634: arm:2R [4092456,4349779]
estimated-cyto:45A7-45C5; 05/16/2001
Plate: RE.693 row: G column: 5
High quality sequence stop: 607.
Location/Qualifiers

FEATURES
source
1..700
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="R669377"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo pF1c-1"
/note="Organ: embryo; Vector: pF1c1; Site 1: XhoI; Site 2: BamHI. Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

BASE COUNT 143 a 202 c 206 g 148 t 1 others

ORIGIN
Query Match 32.4%; Score 603; DB 12; Length 700;
Best Local Similarity 99.8%; Pred. No. 3.9e-114;
Matches 603; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 157 GTAAGGAATCCAGCGGATGTTGGAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 216
DB 97 GTAAGGAATCCAGCGGATGTTGGAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156
QY 217 GCCGTGGCGAGCAAGCGCCGCTTTTGTGGGAGCCCAAGTGTGGGAAACCTTC 276
DB 157 GCCGTGGCGAGCAAGCGCCGCTTTTGTGGGAGCCCAAGTGTGGGAAACCTTC 216
QY 277 CTGAAGAGCGGTGCTCCAGTGAAGTTGCCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 336
DB 217 CTGAAGAGCGGTGCTCCAGTGAAGTTGCCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
QY 337 ATGCTGCTGCTTTCAGAGAAATGCTGAGAGCAAGCACTTCCTGCTGCTCACTCC 396
DB 277 ATGCTGCTGCTTTCAGAGAAATGCTGAGAGCAAGCACTTCCTGCTGCTCACTCC 336

QY 397 CAGGCCAGTCTGCTACGCCGAGTGAGTGAAGCCCAAGACTTACTACACAGC 456
|
|
|
Db 337 CAGGCCAGTCTGCTACGCCGAGTGAGTGAAGCCCAAGACTTACTACACAGC 396
|
|
|
QY 457 GTGGAGAACCCCTCGAGGAGTGCCTGAGTGGCCCAAGGCGAGCAACTCTCAGC 516
|
|
|
Db 397 GTGGAGAACCCCTCGAGGAGTGCCTGAGTGGCCCAAGGCGAGCAACTCTCAGC 456
|
|
|
QY 517 GATGCAAGTGAAGTGAACCAAGCCGAGTGTGCTAGTGAAGTGTGCTGTTTGTG 576
|
|
|
Db 457 GATGCAAGTGAAGTGAACCAAGCCGAGTGTGCTAGTGAAGTGTGCTGTTTGTG 516
|
|
|
QY 577 ACCTTGAAGAGCCGCGGAGAGCCGAGGAGTGTGAGTGCATGAGCCGCGCAT 636
|
|
|
Db 517 ACCTTGAAGAGCCGCGGAGAGCCGAGGAGTGTGAGTGCATGAGCCGCGCAT 576
|
|
|
QY 637 GCTGCTATCAGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 636
|
|
|
Db 577 GCTGCTATCAGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 636
|
|
|
QY 697 ACGGCTCTGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 756
|
|
|
Db 637 ACGGCTCTGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 696
|
|
|
QY 757 ATGT 760
|
|
|
Db 697 ATGT 700

RESULT 12
AA391125 714 bp mRNA linear EST 19-APR-2001
LOCUS LD10222.5prime LD Drosophila melanogaster embryo Bluescript
DEFINITION Drosophila melanogaster cDNA clone LD10222 5prime, mRNA sequence.
ACCESSION AA391125
VERSION AA391125.1 GI:2790665
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 714)
AUTHORS Harvey D., Brokstein P., Hong L., Evans-Holm M., Su C., Tsang G.,
Lewis S. and Rubin G.M.
TITLE BDGP/HMI Drosophila EST Project
JOURNAL Unpublished
COMMENT On Jan 17, 1998 this sequence version replaced gi:2044331.
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Plate: 102 row: B column: 10
High quality sequence stop: 626.

FEATURES

source

1. 714

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="BDGP EST:BDCLN009489"

/db_xref="taxon:7222"

/clone="LD10222"

/sex="male and female"

/dev_stage="0 to 24 hours mixed stage embryonic"

/lab_host="SOLR"

/note="Organ: embryo; Vector: Bluescript SK; Site: 1: EcoRI
; Site 2: XhoI; Constructed using Stratagene ZAP-CDNA
Synthesis Kit. Oligo dT-primed and directionally cloned at
EcoRI and XhoI in Bluescript SK(+/-)"

BASE COUNT

148 a 211 c 204 g 149 t 2 others

Query Match 32.1%; Score 598.2; DB 9; Length 714;
Best Local Similarity 97.8%; Pred. No. 3.8e-113;
Matches 637; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

QY 157 GTTAGCAATCCAGCCGATGTTGTGGAATGCTGATTTGCTGCTGATTTGGGCT 216
|
|
|
Db 68 GTTAGCAATCCAGCCGATGTTGTGGAATGCTGATTTGCTGCTGATTTGGGCT 127
|
|
|
QY 217 GCCGTGGGAGAGAAAGCCGCTTTTGTGGGAGCAAGTGTGGAAAGCCCTCC 276
|
|
|
Db 128 GCCGTG--GGAAGAAAGCCGCTTTTGTGGGAGCAAGTGTGGAAAGCCCTCC 185
|
|
|
QY 277 CTGAAGACGATGCTCCCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 336
|
|
|
Db 186 CTGAAGACGATGCTCCCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 245
|
|
|
QY 337 ATGTGCTGAGCTTTCAGAGAAATGAGCTGAGAGCAAGACTTCTGTGCTCAATCC 396
|
|
|
Db 246 ATGTGCTGAGCTTTCAGAGAAATGAGCTGAGAGCAAGACTTCTGTGCTCAATCC 305
|
|
|
QY 397 CAGGCGAGTCTGCTACAGCCGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 456
|
|
|
Db 306 CAGGCGAGTCTGCTACAGCCGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 365
|
|
|
QY 457 GTGGAGAACCCCTCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 516
|
|
|
Db 366 GTGGAGAACCCCTCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 425
|
|
|
QY 517 GATGCAAGTGAAGTGAACCAAGCCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 576
|
|
|
Db 426 GATGCAAGTGAAGTGAACCAAGCCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 484
|
|
|
QY 577 ACCTTGAAGAGAGCCGCGGAGAGCCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 636
|
|
|
Db 485 ACCTTGAAGAGAGCCGCGGAGAGCCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 544
|
|
|
QY 637 GCTGCTATCAGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 696
|
|
|
Db 545 GCTGCTATCAGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 603
|
|
|
QY 697 ACGGCTCTGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 756
|
|
|
Db 604 ACGGCTCTGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 663
|
|
|
QY 757 ATGTGAGATGCAACATGATTTGATTTCTTACATGCTGCTGCTTAC 807
|
|
|
Db 664 ATGTGAGATGCAACATGATTTGATTTCTTACATGCTGCTGCTTAC 714

RESULT 13
BI622312 596 bp mRNA linear EST 07-SEP-2001
LOCUS RH54158.5prime RH Drosophila melanogaster normalized Head pR1C-1
DEFINITION Drosophila melanogaster cDNA clone RH54158 5 similar to CG8029:
PBan0008029 GO: [enzyme (GO:0003824)] located on: 2R 45B7-45B7.1;
08/22/2001, mRNA sequence.

ACCESSION BI622312
VERSION BI622312.1 GI:15517837
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 596)
AUTHORS Stapleton M., Brokstein P., Hong L., Tyler D., Berman B., Carlson
J., Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George
R., Gonzalez M., Guarin H., Harris N., Li P., Liao G., Mistr S.,
Mungall C.J., Nuno J., Pacled J., Paragas V., Park S.,
Phouanavong S., Wan K., Yu C., Lewis S.E., Ceiniker S. and Rubin
G.M.

BDGP/HMI RH Drosophila EST Project

QY 241 TTTTGTGGGAGCCAAAGTGTGGGAAACCTCTCCCTGAAGACGGTGTCCCAAGTGAG 300
AA141026/c
Db 253 TTTTGTGGGAGCCAAAGTGTGGGAAACCTCTCCCTGAAGACGGTGTCCCAAGTGAG 312
QY 301 TTTTGTGGGAGCCAAAGTGTGGGAAACCTCTCCCTGAAGACGGTGTCCCAAGTGAG 360
Db 313 TTTTGTGGGAGCCAAAGTGTGGGAAACCTCTCCCTGAAGACGGTGTCCCAAGTGAG 372
QY 361 GGCCTGAGAGCAAGGACTTCTGTGTCTCAACTCCAGGCGAGTCTGTACGCCAG 420
Db 373 GGCCTGAGAGCAAGGACTTCTGTGTCTCAACTCCAGGCGAGTCTGTACGCCAG 432
QY 421 CTGCAAGGAGTGAAGCCCAAGACTTAACCAAGGCGAGGAGAACCCCTCGAGGACATG 480
Db 433 CTGCAAGGAGTGAAGCCCAAGACTTAACCAAGGCGAGGAGAACCCCTCGAGGACATG 492
QY 481 CGCTCAGTGGCCGCAAGGCGAGGCAAACTCCATGATGCCAGTGAAGAGCTGACCAAG 540
Db 493 CGCTCAGTGGCCGCAAGGCGAGGCAAACTCCATGATGCCAGTGAAGAGCTGACCAAG 552
QY 541 CGGCGCAAGTGGCCGAGTGAAGGCGGCTGTGTGTGACCTTGAGAGGAGCGCGCG 595
Db 553 CGGCGCAAGTGGCCGAGTGAAGGCGGCTGTGTGTGACCTTGAGAGGAGCGCGCG 607

RESULT 15

AA141026 592 bp mRNA linear EST 29-NOV-1998

LOCUS CK01297.3prime CK Drosophila melanogaster embryo Bluescript

DEFINITION Drosophila melanogaster cDNA clone CK01297 3prime, mRNA sequence.

ACCESSION AA141026

VERSION AA141026.1 GI:1704633

KEYWORDS EST.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 592)

Kopczynski, C., Serano, T., Rubin, G. and Goodman, C.

BDGP/HMI CK Drosophila EST Project

Unpublished

Other ESTs: CK01297.5prime

COMMENT Contact: Stapleton, M.

BDGP Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu

Insert Length: 592 Std Error: 0.00

Plate: 10 row: B column: 12

High quality sequence stop: 592.

Location/Qualifiers

1. 592

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="BDGP_EST:BDcln000653"

/db_xref="taxon:7227"

/clone="CK01297"

/sex="male and female"

/dev_stage="0 to 24 hours old embryo"

/lab_host="XLI Blue MRP"

/clone_lib="CK Drosophila melanogaster embryo Bluescript"

/note="Organ: Vector: Bluescript SK; Site 1: Clat;

Site 2: PstI; mRNA purified from rough endoplasmic

reticulum-bound polysomes was used as a template. cDNA's

directionally cloned at HindIII and PstI in Bluescript

SK. Cloned into HindIII, but the site was destroyed to

add an adapter sequence. So Clat is now the restriction

enzyme at site 1 of vector. Primers - 5' universal, 3'

m3-20 (reverse), 3' T3, 5' T7."

BASE COUNT 189 a 112 c 133 g 157 t 1 others

Query Match 31.5%; Score 587.4; DB 9; Length 592;
Best Local Similarity 99.7%; Pred. No. 6.3e-111;
Matches 588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1274 CCAAGGGCAAGACTATCAACATCAATTCGCCGCCGAGTAAACCGCCATCATTCGGTT 1333
Db 592 CCAGGGGCAAGACTATCAACATCAATTCGCCGCCGAGTAAACCGCCATCATTCGGTT 533
QY 1334 CACACCTTATTGATTGTTCAATGGCACTGAATATTCCTTATTCCTTCACTGTA 1393
Db 532 CACACCTTATTGATTGTTCAATGGCACTGAATATTCCTTATTCCTTCACTGTA 473
QY 1394 TCACATGATATATGTTAAAGATCAATCTTAATACGATACAGGATTAACATTAACA 1453
Db 472 TCACATGATATATGTTAAAGATCAATCTTAATACGATACAGGATTAACATTAACA 413
QY 1454 TAACCGGGAATTAATGCTCTCAATTTGTATGTCGTAATTTTCCCGCAAGATC 1513
Db 412 TAACCGGGAATTAATGCTCTCAATTTGTATGTCGTAATTTTCCCGCAAGATC 353
QY 1514 AGTTTACGTTGAGATCAAGGCTGCTGATTTCTGTGCGCGCAAAATTCCTGAGTTT 1573
Db 352 AGTTTACGTTGAGATCAAGGCTGCTGATTTCTGTGCGCGCAAAATTCCTGAGTTT 293
QY 1574 TGCTTGCAACGAGCAGGATCCAGTATTTCAATTAAGATAGCCCTTAATGCTTTATC 1633
Db 292 TGCTTGCAACGAGCAGGATCCAGTATTTCAATTAAGATAGCCCTTAATGCTTTATC 233
QY 1634 GAAGAACTTTAAAGACGACATGTGTGTGTGTTAAGCGCGCGGTGTGTGT 1693
Db 232 GAAGAACTTTAAAGACGACATGTGTGTGTGTTAAGCGCGCGGTGTGTGTGT 173
QY 1694 GTGCAATGGGAGTAATCTTAATCCGAATATCTAAGTGTATTAAGTTCTTGT 1753
Db 172 GTGCAATGGGAGTAATCTTAATCCGAATATCTAAGTGTATTAAGTTCTTGT 113
QY 1754 GCGCAATATGCGAGACTAATTTATGCTGTATGATGATGTTGATTTGCAACCC 1813
Db 112 GCGCAATATGCGAGACTAATTTATGCTGTATGATGATGTTGATTTGCAACCC 53
QY 1814 TCGTGTGTGCGAGGCGCCCAAGAAATTAAGTTTTCATTTGTTA 1863
Db 52 TCGTGTGTGCGAGGCGCCCAAGAAATTAAGTTTTCATTTGTTA 3

Search completed: February 8, 2004, 10:22:06
Job time : 4377.39 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:35:30 ; Search time 9744.18 Seconds
(without alignments)
10739.422 Million cell updates/sec

Title: US-09-614-150A-8

Perfect score: 2558

Sequence: 1 ggcgcgcgccttccttc.....taatacaaatcaagaacaa 2558

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481336 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBankl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rnd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_hcg_hum:*
39: em_hcg_mus:*
40: em_hcg_mus:*
41: em_hcg_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2499.2	97.7	2525	3	AY121641	AY121641 Drosophil
2	2436.4	95.2	2563	3	DMRHO	X52454 D.melanog
3	1723.4	67.4	55700	2	AC017867	AC017867 Drosophil
4	1723.4	67.4	196594	2	AC010564	AC010564 Drosophil
5	1723.4	67.4	302131	3	AE003471	AE003471 Drosophil
6	795.2	31.1	49858	2	AC008252	AC008252 Drosophil
7	597	23.3	1797	2	AB089248	AB089248 Drosophil
8	251.8	9.8	1816	3	AY118997	AY118997 Drosophil
9	251.8	9.8	1822	3	AF318284	AF318284 Drosophil
10	198.2	7.7	2442	3	AF318283	AF318283 Drosophil
11	197	7.7	1900	3	AY119150	AY119150 Drosophil
12	197	7.7	2107	3	BT003499	BT003499 Drosophil
13	170.8	6.7	35808	2	AC017712	AC017712 Drosophil
14	169.8	6.6	175481	3	AC010007	AC010007 Drosophil
15	169.8	6.6	268369	3	AC005847	AC005847 Drosophil
16	162.2	6.3	199	3	AF067878	AF067878 Drosophil
17	161.4	6.3	1422	9	HSB272344	AX375249 Sequence
18	161.4	6.3	1559	6	HSY17108	AX375249 Sequence
19	161.4	6.3	1559	9	AF067882	Y17108 Homo sapien
20	160	6.3	204	3	AF067882	AF067882 Drosophil
21	160	6.3	1376	6	AX375253	AX375253 Sequence
22	160	6.3	1419	9	HSB313480	AJ313480 Homo sapi
23	159.8	6.2	195	3	AF067877	AF067877 Drosophil
24	158.4	6.2	198	3	AF067881	AF067881 Drosophil
25	156.8	6.1	198	3	AF067879	AF067879 Drosophil
26	156.4	6.1	1557	10	BC021549	BC021549 Mus muscu
27	155	6.1	33554	3	AC004343	AC004343 Drosophil
28	155	6.1	55700	2	AC017867	AC017867 Drosophil
29	155	6.1	196594	3	AC010564	AC010564 Drosophil
30	155	6.1	302131	3	AE003471	AE003471 Drosophil
31	154.6	6.0	203	3	AF067880	AF067880 Drosophil
32	148.2	5.8	1320	6	AX600881	AX600881 Sequence
33	145.6	5.7	2049	3	AF318285	AF318285 Drosophil
34	145.6	5.7	2389	3	AY061605	AY061605 Drosophil
35	143.2	5.6	2022	10	NM0313479	AJ313479 Mus muscu
36	143.2	4.4	493	10	RNY17258	Y17258 Rattus norv
37	109.4	4.3	66993	2	AC138074	AC138074 Homo sapi
38	108.2	4.2	34033	2	AC017819	AC017819 Drosophil
39	108.2	4.2	155000	2	AC023689	AC023689 Drosophil
40	108.2	4.2	328128	3	AE003486	AE003486 Drosophil
41	103.4	4.0	93791	2	AC138073	AC138073 Homo sapi
42	101	3.9	170627	2	AC125567	AC125567 Rattus no
43	97.2	3.8	76245	2	AC125151_4	Continuation (5 of
44	97.2	3.8	131874	2	AC131705	AC131705 Mus muscu
45	96.8	3.8	349980	6	AX344570	AX344570 Sequence

ALIGNMENTS

RESULT 1
LOCUS AY121641 2525 bp mRNA linear INV 18-JUN-2002
DEFINITION Drosophila melanogaster LD06131 full insert cDNA.
ACCESSION AY121641
VERSION AY121641.1 GI:21464329
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
Stapleton M., Brokstein P., Hong D., Agbayani A., Carlson J.,
1 (bases 1 to 2525)
Champe M., Chavez C., Dorsett V., Dreenek D., Farfan D., Friese E.,

Pred. No. is the number of results predicted by chance to have a

OY	1364	AGCTATGAGAGAGAGGAGTGAGTGATTAACCAAGACTCTGTGGACAACCTGGAGAGTG	1423
Db	1331	AGCTATGAGAGAGAGGAGTGAGTGATTAACCAAGACTCTGTGGACAACCTGGAGAGTG	1380
OY	1424	CCTAAGTTGAGTTCGAGTTCGTGAGCATGCTCGACAGGATTCGGAACTCTGTTGAGCT	1483
Db	1381	CCTAAGTTTGAAGTTTCGAGTTCGTGAGCATGCTCGACAGGATTCGGAACTCTGTTGAGCT	1440
OY	1484	TCAGGAGAGATTCGAGAGACAGAGAGTTGGTGAAGAAAGAAAGTTCACTCAACGATTAGT	1543
Db	1441	TCAGGAGAGATTCGAGAGACAGAGAGTTGGTGAAGAAAGAAAGTTCACTCAACGATTAGT	1500
OY	1544	TCAAACCTAATTCGATTCGTTTCGTTGGCTTTGCTTTGTTAGCATATATCTCGATTACGT	1603
Db	1501	TCAAACCTAATTCGATTCGTTTCGTTGGCTTTGCTTTGTTAGCATATATCTCGATTACGT	1560
OY	1604	TACCGTTTGCAAGTTTAAAGCTTTCAGTTCCGAAACATAGTATACAAACTCAAAAAAAA	1663
Db	1561	TACCGTTTGCAAGTTTAAAGCTTTCAGTTCCGAAACATAGTATACAAACTCAAAAAAAA	1620
OY	1664	AAACAAATCAAGAGAAATACACTGGA-CAAAAAAGAAAGTCGAGAGTGAAGAGAACTA	1722
Db	1621	AAACAAATCAAGAGAAATACACTGGA-CAAAAAAGAAAGTCGAGAGTGAAGAGAACTA	1680
OY	1723	AACCGAACCCGAAAGCTGTAAACAAATGTGTGATAGAACCAAGAACTGAATTTATTTGCG	1782
Db	1681	AACCGAACCCGAAAGCTGTAAACAAATGTGTGATAGAACCAAGAACTGAATTTATTTGCG	1740
OY	1783	CGTGAAACCAAGTAAACAAATCAAGAGAAATCAAGAGAGAGAAACAGAACTAATTCGC	1842
Db	1741	CGTGAAACCAAGTAAACAAATCAAGAGAAATCAAGAGAGAGAAACAGAACTAATTCGC	1800
OY	1843	CTCTGCTATGATTTAAATGATTTATCCAGTGTTCAAATTAATGTTTGTCTGTTTCTT	1902
Db	1801	CTCTGCTATGATTTAAATGATTTATCCAGTGTTCAAATTAATGTTTGTCTGTTTCTT	1860
OY	1903	AAATTAATGATTTATTTGGCCGCAATTACTAGAAATGCAATCGAATCGAACATCAGCAAC	1962
Db	1861	AAATTAATGATTTATTTGGCCGCAATTACTAGAAATGCAATCGAATCGAACATCAGCAAC	1920
OY	1963	TGTAATCAATTTGTTATATCATCCATPACATPATATGTCGCCAATTTGAGAGTTTAGTGT	2022
Db	1921	TGTAATCAATTTGTTATATCATCCATPACATPATATGTCGCCAATTTGAGAGTTTAGTGT	1980
OY	2023	TATATTTATATATTTAGGTATTAAGTATACAGCTCTCTTAACAAATTTGTTCAATTTGTAA	2082
Db	1981	TATATTTATATATTTAGGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA	2040
OY	2083	TACTATTAAGTCGACACTAGTCMAAACACACACACACACACACACACACACACACAAAA	2142
Db	2041	TACTATTAAGTCGACACTAGTCMAAACACACACACACACACACACACACACACACAAAA	2100
OY	2143	AATATGAAACACACAGCAAGAACACATTCATTCAGATCAATTTAAGCAATTCGAGTTAA	2202
Db	2101	AATATGAAACACACAGCAAGAACACATTCATTCAGATCAATTTAAGCAATTCGAGTTAA	2160
OY	2203	ATTAAATTAATTTACTAAAGTCACTTAATGCGTTTCAAAAATCGAGCATGTAATCCCC	2262
Db	2161	ATTAAATTAATTTACTAAAGTCACTTAATGCGTTTCAAAAATCGAGCATGTAATCCCC	2220
OY	2263	TACACACACACACACACACACACTCGAAAGATTAATTAATTTATTTATTTATGTTAG	2322
Db	2221	TACACACACACACACACACACACTCGAAAGATTAATTAATTTATTTATTTATTTATGTTAG	2280
OY	2323	GGCAGCGAGGTTTATTAATTCGTCAATTTGAGCGAACTAATTTATTTATTTATTTTAA	2382
Db	2281	GGCAGCGAGGTTTATTAATTCGTCAATTTGAGCGAACTAATTTATTTATTTATTTTAA	2340
OY	2383	TAAATTTAGTAAATTCACACAAACAGACAGAAAAACACACACACACAGAGAGAGAA	2442
Db	2341	TAAATTTAGTAAATTCACACAAACAGACAGAAAAACACACACACACAGAGAGAGAA	2400
OY	2443	GAACAAACCAATTCACCTGTAAAAATATCCAAATTTGAAAAATACACAGAAAGCCAAAG	2502

Db	2401	GAACCAACCAATTCACACTGTAAAAAATTCCAATTGAAAAATACACGAAACCCAAAG	2460
Oy	2503	AAATTAATAATCAAAATTCATTCAGAAATACACAGTAATACAAATACAGCAA	2558
Db	2461	AAATTAATAATCAAAATTCATTCAGAAATACACAGTAATACAAATTAATAATAA	2516
RESULT 2			
DMRHO			
LOCUS	D. melanogaster rho gene.	2563 bp	linear
DEFINITION			INV 25-FEB-1993
ACCESSION	X52454		
VERSION	X52454.1	GI:10878	
KEYWORDS	rho gene.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
	Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 2563)		
AUTHORS	Bier, E., Jan, L.Y. and Jan, Y.N.		
TITLE	rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous system development in Drosophila melanogaster		
JOURNAL	Genes Dev. 4 (2), 190-203 (1990)		
MEDLINE	90249726		
PUBMED	2110920		
FEATURES			
source	location/Qualifiers		
	1..2563		
	/organism="Drosophila melanogaster"		
	/mol_type="mRNA"		
	/db_xref="taxon:7227"		
gene	1..2563		
	/gene="rho"		
CDS	361..1428		
	/gene="rho"		
	/codon_start=1		
	/protein_id="CAA36692.1"		
	/db_xref="GI:10879"		
	/db_xref="FLYBASE:FBgn0004635"		
	/db_xref="SWISS-PROT:P20350"		
	/translation="MEMBRANOUS PROTEIN; DROSOPHILA MELANOGASTER; RHO; RHO-1; RHO-2; RHO-3; RHO-4; RHO-5; RHO-6; RHO-7; RHO-8; RHO-9; RHO-10; RHO-11; RHO-12; RHO-13; RHO-14; RHO-15; RHO-16; RHO-17; RHO-18; RHO-19; RHO-20; RHO-21; RHO-22; RHO-23; RHO-24; RHO-25; RHO-26; RHO-27; RHO-28; RHO-29; RHO-30; RHO-31; RHO-32; RHO-33; RHO-34; RHO-35; RHO-36; RHO-37; RHO-38; RHO-39; RHO-40; RHO-41; RHO-42; RHO-43; RHO-44; RHO-45; RHO-46; RHO-47; RHO-48; RHO-49; RHO-50; RHO-51; RHO-52; RHO-53; RHO-54; RHO-55; RHO-56; RHO-57; RHO-58; RHO-59; RHO-60; RHO-61; RHO-62; RHO-63; RHO-64; RHO-65; RHO-66; RHO-67; RHO-68; RHO-69; RHO-70; RHO-71; RHO-72; RHO-73; RHO-74; RHO-75; RHO-76; RHO-77; RHO-78; RHO-79; RHO-80; RHO-81; RHO-82; RHO-83; RHO-84; RHO-85; RHO-86; RHO-87; RHO-88; RHO-89; RHO-90; RHO-91; RHO-92; RHO-93; RHO-94; RHO-95; RHO-96; RHO-97; RHO-98; RHO-99; RHO-100; RHO-101; RHO-102; RHO-103; RHO-104; RHO-105; RHO-106; RHO-107; RHO-108; RHO-109; RHO-110; RHO-111; RHO-112; RHO-113; RHO-114; RHO-115; RHO-116; RHO-117; RHO-118; RHO-119; RHO-120; RHO-121; RHO-122; RHO-123; RHO-124; RHO-125; RHO-126; RHO-127; RHO-128; RHO-129; RHO-130; RHO-131; RHO-132; RHO-133; RHO-134; RHO-135; RHO-136; RHO-137; RHO-138; RHO-139; RHO-140; RHO-141; RHO-142; RHO-143; RHO-144; RHO-145; RHO-146; RHO-147; RHO-148; RHO-149; RHO-150; RHO-151; RHO-152; RHO-153; RHO-154; RHO-155; RHO-156; RHO-157; RHO-158; RHO-159; RHO-160; RHO-161; RHO-162; RHO-163; RHO-164; RHO-165; RHO-166; RHO-167; RHO-168; RHO-169; RHO-170; RHO-171; RHO-172; RHO-173; RHO-174; RHO-175; RHO-176; RHO-177; RHO-178; RHO-179; RHO-180; RHO-181; RHO-182; RHO-183; RHO-184; RHO-185; RHO-186; RHO-187; RHO-188; RHO-189; RHO-190; RHO-191; RHO-192; RHO-193; RHO-194; RHO-195; RHO-196; RHO-197; RHO-198; RHO-199; RHO-200; RHO-201; RHO-202; RHO-203; RHO-204; RHO-205; RHO-206; RHO-207; RHO-208; RHO-209; RHO-210; RHO-211; RHO-212; RHO-213; RHO-214; RHO-215; RHO-216; RHO-217; RHO-218; RHO-219; RHO-220; RHO-221; RHO-222; RHO-223; RHO-224; RHO-225; RHO-226; RHO-227; RHO-228; RHO-229; RHO-230; RHO-231; RHO-232; RHO-233; RHO-234; RHO-235; RHO-236; RHO-237; RHO-238; RHO-239; RHO-240; RHO-241; RHO-242; RHO-243; RHO-244; RHO-245; RHO-246; RHO-247; RHO-248; RHO-249; RHO-250; RHO-251; RHO-252; RHO-253; RHO-254; RHO-255; RHO-256; RHO-257; RHO-258; RHO-259; RHO-260; RHO-261; RHO-262; RHO-263; RHO-264; RHO-265; RHO-266; RHO-267; RHO-268; RHO-269; RHO-270; RHO-271; RHO-272; RHO-273; RHO-274; RHO-275; RHO-276; RHO-277; RHO-278; RHO-279; RHO-280; RHO-281; RHO-282; RHO-283; RHO-284; RHO-285; RHO-286; RHO-287; RHO-288; RHO-289; RHO-290; RHO-291; RHO-292; RHO-293; RHO-294; RHO-295; RHO-296; RHO-297; RHO-298; RHO-299; RHO-300; RHO-301; RHO-302; RHO-303; RHO-304; RHO-305; RHO-306; RHO-307; RHO-308; RHO-309; RHO-310; RHO-311; RHO-312; RHO-313; RHO-314; RHO-315; RHO-316; RHO-317; RHO-318; RHO-319; RHO-320; RHO-321; RHO-322; RHO-323; RHO-324; RHO-325; RHO-326; RHO-327; RHO-328; RHO-329; RHO-330; RHO-331; RHO-332; RHO-333; RHO-334; RHO-335; RHO-336; RHO-337; RHO-338; RHO-339; RHO-340; RHO-341; RHO-342; RHO-343; RHO-344; RHO-345; RHO-346; RHO-347; RHO-348; RHO-349; RHO-350; RHO-351; RHO-352; RHO-353; RHO-354; RHO-355; RHO-356; RHO-357; RHO-358; RHO-359; RHO-360; RHO-361; RHO-362; RHO-363; RHO-364; RHO-365; RHO-366; RHO-367; RHO-368; RHO-369; RHO-370; RHO-371; RHO-372; RHO-373; RHO-374; RHO-375; RHO-376; RHO-377; RHO-378; RHO-379; RHO-380; RHO-381; RHO-382; RHO-383; RHO-384; RHO-385; RHO-386; RHO-387; RHO-388; RHO-389; RHO-390; RHO-391; RHO-392; RHO-393; RHO-394; RHO-395; RHO-396; RHO-397; RHO-398; RHO-399; RHO-400; RHO-401; RHO-402; RHO-403; RHO-404; RHO-405; RHO-4		

Db 241 GAAAAAGTGCAGATATGCCAACTAAAGACATTGGATTACAAGAAACCCAGCAT 300
Qy 301 TTGGATTATTAACATTGGCGACAGGCGAAAAAACCCTAAGATTCTTCAACGGGCGCAGC 360
Db 301 TTGGATTATTAACATTGGCGACAGGCGAAAAAACCCTAAGATTCTTCAACGGGCGCAGC 360
Qy 361 ATGAGAACTTAAACGCAAAATGTAAACGAAACCAAGGTGAAATTTGGGCGAGAGAAAGAG 420
Db 361 ATGAGAAACCCAAACGCAAAATGTAAACGAAACCAAGGTGAAATTTGGGCGAGAGAAAGAG 420
Qy 421 AAGAGAGGCTTCGAG 480
Db 421 AAGAGAGGCTTCGAG 480
Qy 481 CCGGCGGCTGCTCAGATTCTCCAACTCCCTGCTGATGAGACACGAAATGAGAGAGAGAG 540
Db 481 CCGGCGGCTGCTCAGATTCTCCAACTCCCTGCTGATGAGACACGAAATGAGAGAGAGAG 540
Qy 541 AGCAGCAGCTGCTGACACCGCGCAAGGAGAGACATCTAATGCAACGAGGCGCATCCG 600
Db 541 AGCAGCAGCTGCTGACACCGCGCAAGGAGAGACATCTAATGCAACGAGGCGCATCCG 600
Qy 601 GCCACGCACTTCCGAGAGTGGAGAGATATGCGCTGCTGAAGTATGCTTCAACCGCAGAC 660
Db 601 GCCACGCACTTCCGAGAGTGGAGAGATATGCGCTGCTGAAGTATGCTTCAACCGCAGAC 660
Qy 661 TGGGCGGCTGCTCAGATTCTCCAACTCCCTGCTGATGAGACACGAAATGAGAGAGAG 720
Db 661 TGGGCGGCTGCTCAGATTCTCCAACTCCCTGCTGATGAGACACGAAATGAGAGAGAG 720
Qy 721 TACACAAATGCTGCTGACACCGCGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 TACACAAATGCTGCTGACACCGCGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 781 GTCTATGCGCGGAGACCGGCGCTGCTGAGGTGGCGCTTCTTAACTACATGTTCTGAC 840
Db 781 GTCTATGCGCGGAGACCGGCGCTGCTGAGGTGGCGCTTCTTAACTACATGTTCTGAC 840
Qy 841 GCCAACTGCTTCCAGCTGAGTCTGATGATGATGATGATGATGATGATGATGATGATG 900
Db 841 GCCAACTGCTTCCAGCTGAGTCTGATGATGATGATGATGATGATGATGATGATGATG 900
Qy 901 GAGGTGATGACAGGCGACAGGCGAGAGATGCGGTGATCTAATGCGGCGGCTTTTGGCGGA 960
Db 901 GAGGTGATGACAGGCGACAGGCGAGAGATGCGGTGATCTAATGCGGCGGCTTTTGGCGGA 960
Qy 961 TCCCTGGGCAACAGTGTGCTGACTCGAGGTCTTCTGCTGAGGCGCGAGCGGTGCGT 1020
Db 961 TCCCTGGGCAACAGTGTGCTGACTCGAGGTCTTCTGCTGAGGCGCGAGCGGTGCGT 1020
Qy 1021 TATGCGCTGTTGGCGGACATCTGGCGAATTAATCACTGAATTAATGCGCATGTAAGAGC 1080
Db 1021 TATGCGCTGTTGGCGGACATCTGGCGAATTAATCACTGAATTAATGCGCATGTAAGAGC 1080
Qy 1081 GCATTCACGCAACTCGGATCGGTGATCTTTGCTCTGCGATCTGCGGTATGCTCTC 1140
Db 1081 GCATTCACGCAACTCGGATCGGTGATCTTTGCTCTGCGATCTGCGGTATGCTCTC 1140
Qy 1141 TACACCAATTAATCTGATGAAAGCGCTTGGCGAAGGCTCCAGGTGTGCTACATTTCC 1200
Db 1141 TACACCAATTAATCTGATGAAAGCGCTTGGCGAAGGCTCCAGGTGTGCTACATTTCC 1200
Qy 1201 CACTGACGGAAGCGCTGGGAGAGACTAATCGGCTTTCTGCTGAAAGAACTTGGGT 1260
Db 1201 CACTGACGGAAGCGCTGGGAGAGACTAATCGGCTTTCTGCTGAAAGAACTTGGGT 1260
Qy 1261 CACGAGAGTACAG 1320
Db 1261 CACGAGAGTACAG 1320
Qy 1321 GTCTTGCATCGTTTCAACCTGATCAACCGGTGACCGGCTGATGAGAGAGAG 1380
Db 1321 GTCTTGCATCGTTTCAACCTGATCAACCGGTGACCGGCTGATGAGAGAGAG 1380

Qy 1381 GTGAGGTGATTAACCGAGATCTGTTGACGAGCCTGGAGGTGCTTAAGTTGAGGTTG 1440
Db 1381 GTGAGGTGATTAACCGAGATCTGTTGACGAGCCTGGAGGTGCTTAAGTTGAGGTTG 1440
Qy 1441 GAGTGTGACAGATCTGCGAGGAGATTCGGAATTCGTTGAGCTTCAGAGAGATTCAGAG 1500
Db 1441 GAGTGTGACAGATCTGCGAGGAGATTCGGAATTCGTTGAGCTTCAGAGAGATTCAGAG 1500
Qy 1501 ACAGAGATGAGGAG 1560
Db 1501 ACAGAGATGAGGAG 1560
Qy 1561 TTGCTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTT 1620
Db 1561 TTGCTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTT 1620
Qy 1621 CGTTTCACTTCGAAACATGATCAACAACTCAAAATCAAAATCAAAATCAAAATCA 1680
Db 1621 CGTTTCACTTCGAAACATGATCAACAACTCAAAATCAAAATCAAAATCAAAATCA 1680
Qy 1681 TACACTGACAAAG 1740
Db 1681 TACACTGACAAAG 1740
Qy 1741 TAAACAAATGCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 1741 TAAACAAATGCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Qy 1801 AAATCAAG 1860
Db 1801 AAATCAAG 1860
Qy 1861 TG-----ATTATCAATGTTTCAATTAATGTTTCAATTAATGTTTCAATTAATGTTT 1915
Db 1861 TG-----ATTATCAATGTTTCAATTAATGTTTCAATTAATGTTTCAATTAATGTTT 1915
Qy 1915 ATTGCGCGCAATTAACAG 1975
Db 1915 ATTGCGCGCAATTAACAG 1975
Qy 1975 TTAATCAATCAAG 2035
Db 1975 TTAATCAATCAAG 2035
Qy 2035 TATTTAGTATTAATGAG 2095
Db 2035 TATTTAGTATTAATGAG 2095
Qy 2095 CACACTAGTCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 2155
Db 2095 CACACTAGTCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 2155
Qy 2155 ACAGCAAG 2215
Db 2155 ACAGCAAG 2215
Qy 2215 ATACTAAGTCACTTAATGAGTCAAAATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2272
Db 2215 ATACTAAGTCACTTAATGAGTCAAAATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2272
Qy 2272 CAC 2332
Db 2272 CAC 2332
Qy 2332 GTTTATTAATTCGTAATGAGGAACTAATTAATTAATTAATTAATTAATTAATTAAT 2392
Db 2332 GTTTATTAATTCGTAATGAGGAACTAATTAATTAATTAATTAATTAATTAATTAAT 2392
Qy 2392 AAATTCACACAAACAG 2452
Db 2392 AAATTCACACAAACAG 2452

QY	2453	AATCAACTGTAAGAAATATCCAAATTGAAATAACAGCGAAAGCCAAACAAATATATATAA	2512
Db	2458	AATTCACCTGTAAGAAATATCCAAATTGAAATAACAGCGAAAGCCAAACAAATATATATAA	2517
QY	2513	ATCAAAACATTTCAAGATACACAGTATATACAAATACAGCAAA 2558	
Db	2518	ATCAAAACATTTCAAGATATACACAGTATATACAAATACAGCAAA 2563	
RESULT 3			
AC017867/c			
LOCUS	AC017867	57700 bp	DNA linear HTG 09-DEC-1999
DEFINITION	Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.		
ACCESSION	AC017867		
VERSION	AC017867.1	GI:5533323	
KEYWORDS	HTG; HTGS; PHASE2.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
AUTHORS	1 (bases 1 to 55700)		
TITLE	Adams, M. and Venter, J. C.		
JOURNAL	Direct Submission		
COMMENT	Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA This sequence was identified as CDW:10212411 by the submitter. For more information on this record e-mail to fly@celera.com. * NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved.		
FEATURES			
source	1..55700		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7227"		
BASE COUNT	14536 a 12616 c 12806 g 15742 t		
ORIGIN			
Query Match	67.4%; Score 1723.4; DB 2; Length 55700; Best Local Similarity 93.3%; Pred. No. 0; Matches 1866; Conservative 0; Mismatches 1; Indels 132; Gaps 1;		
QY	692	TTGAGATTGCATCTTCGCTCAGACCGCTACACAAATGCCGCCAGAAATTTGGGCTAC	751
Db	19257	TTTGAATTTGCATCTTCGCTCAGACCGCTACACAAATGCCGCCAGAAATTTGGGCTAC	19198
QY	752	CCGTTCCGATTCGCTCGGATTCGGTGTCTATTCGGCGAACCGGCGTCTCAGGTGT	811
Db	19197	CCGTTCCGATTCGCTCGGATTCGGTGTCTATTCGGCGAACCGGCGTCTCAGGTGT	19138
QY	812	GGCGCTCTTAGCTACATGTTCTGTGACGCAACGTGTTCCACTGGGCTTCAATATG	871
Db	19137	GGCGCTCTTAGCTACATGTTCTGTGACGCAACGTGTTCCACTGGGCTTCAATATG	19078
QY	872	TCATTCAGCTGTTTATTTGGCAATTCCTCCGTGAGGTGATGACGCGACGGCCAGGATCGCG	931
Db	19077	TCATTCAGCTGTTTATTTGGCAATTCCTCCGTGAGGTGATGACGCGACGGCCAGGATCGCG	19018
QY	932	TGATCTACATGGCGGGCGTTTTTTCGCGAATTCCTCGGACCAAGTGTGACTCGAGG	991
Db	19017	TGATCTACATGGCGGGCGTTTTTTCGCGAATTCCTCGGACCAAGTGTGACTCGAGG	18958
QY	992	TCTTTCCTGGTGGGCGCGAGCGGTGCTATATGCCCTTTGGCCGACATCTGGCCAAAC	1051
Db	18957	TCTTTCCTGGTGGGCGCGAGCGGTGCTATATGCCCTTTGGCCGACATCTGGCCAAAC	18896
QY	1052	TTAACAAGTAATATGCGACATGGAAGAGCGATCAGCAACTCGATCGGTGTATCT	1111
Db	18897	TTAACAAGTAATATGCGACATGGAAGAGCGATCAGCAACTCGATCGGTGTATCT	18833
QY	1112	TT-----	1113

Db	18837	TTGGTAGGTTTCCCAACCTTATATGACTAGAAATCTATAGAACCTATCTATCTTACCTTA	18778
QY	1114	-----	1113
Db	18777	ACAAGATTGTGACATCTTTTCTTCATATGTTTAACCTTAACCTAATTGAATCCCTCTG	18718
QY	1114	-----GTCCTCGGAGATCTGGGCTATATCTCTCTTAACCCCAATCTTCGATG	1159
Db	18717	TTCTTCATTGGAAGTCCTCCGGAATCGGGCTATATCTCTCTTAACCCCAATCTTCGATG	18658
QY	1160	GAAGCGCCTTCGCGCAAGAGGTCGCCAGAGTGTGTGACATATGCCACCTGAACGAGACCTCTG	1219
Db	18657	GAAGCGCCTTCGCGCAAGAGGTCGCCAGAGTGTGTGACATATGCCACCTGAACGAGACCTCTG	18598
QY	1220	CAGGACTAACGATCGGCTTCTGTGTCTTAAGAACTTCCGTCAACGAGATACGACACG	1279
Db	18597	CAGGACTAACGATCGGCTTCTGTGTCTTAAGAACTTCCGTCAACGAGATACGACACG	18538
QY	1280	TCATCTGTGTGCTTAGGCTTGTGGGCTCTACCTGTGCTTCAACCGTCTTCGCAATCGTTTCA	1339
Db	18537	TCATCTGTGTGCTTAGGCTTGTGGGCTCTACCTGTGCTTCAACCGTCTTCGCAATCGTTTCA	18478
QY	1340	ACCGATCAACACGGTGACCGCTCAGCTGATGGAGGACGAGGGTGGATTTACCCACG	1399
Db	18477	ACCGATCAACACGGTGACCGCTCAGCTGATGGAGGACGAGGGTGGATTTACCCACG	18418
QY	1400	ATCTGTGCAACGACCTGGAGTGTCTCTAAGTTTGAGGTTCCGAGTGTCTCAGATCTCGC	1459
Db	18417	ATCTGTGCAACGACCTGGAGTGTCTCTAAGTTTGAGGTTCCGAGTGTCTCAGATCTCGC	18358
QY	1460	AGGGAATCGGAATCTGCTTGAGCTTCAAGAGATCGAGAGACAGAGATGTGTGAGAA	1519
Db	18357	AGGGAATCGGAATCTGCTTGAGCTTCAAGAGATCGAGAGACAGAGATGTGTGAGAA	18298
QY	1520	GAAAAGTCACTCAACGATTTACTTACCTAATAATTCGATTCGTTGGCTTTCCTT	1579
Db	18297	GAAAAGTCACTCAACGATTTACTTACCTAATAATTCGATTCGTTGGCTTTCCTT	18238
QY	1580	TGCTTACCTATATCTCGTTATCGTTACCGTTTGCACTTAAAGTTTCAAGTTCCGAAACAT	1639
Db	18237	TGCTTACCTATATCTCGTTATCGTTACCGTTTGCACTTAAAGTTTCAAGTTCCGAAACAT	18178
QY	1640	AGTACACAACTCAAAAAAAAAAAAAAAAACAATCAGAGAAATCACTGACAAAAAGAA	1699
Db	18177	AGTACACAACTCAAAAAAAAAAAAAAAAACAATCAGAGAAATCACTGACAAAAAGAA	18118
QY	1700	GTGAGAGAGATGAGAGAAACATTAACCCGAGCCGAAACGTGTAAACAAATGTTTGATAG	1759
Db	18117	GTGAGAGAGATGAGAGAAACATTAACCCGAGCCGAAACGTGTAAACAAATGTTTGATAG	18058
QY	1760	AACCAAGACTGAATTTATTTGCGGTGTAAAAACAAGTAAAAATCAAGAGAAAAATCA	1819
Db	18057	AACCAAGACTGAATTTATTTGCGGTGTAAAAACAAGTAAAAATCAAGAGAAAAATCA	17998
QY	1820	AGAGAGAAAACAGAACTAATCGCCTCTCGCTATGATTTAAATGATTTATCATGTTTTCA	1879
Db	17997	AGAGAGAAAACAGAACTAATCGCCTCTCGCTATGATTTAAATGATTTATCATGTTTTCA	17938
QY	1880	ATTATATGTTTGTCTGTTTTCTTAAATTTATGTTATTTATGGCCGCAATTAACGAATCA	1939
Db	17937	ATTATATGTTTGTCTGTTTTCTTAAATTTATGTTATTTATGGCCGCAATTAACGAATCA	17878
QY	1940	ATCGAATCGAAGATCAGCAAACTGTATCAAAATGTTTATACATCAATAGAGATATGTG	1999
Db	17877	ATCGAATCGAAGATCAGCAAACTGTATCAAAATGTTTATACATCAATAGAGATATGTG	17818
QY	2000	CTCCGAATTAGAGATTTAGTGTATATTTATATATTTATAGATATACTAGCCCTCT	2059
Db	17817	CTCCGAATTTAGAGATTTAGTGTATATATTTATATATTTAGATATACTAGCCCTCT	17758
QY	2060	AACAAATGTTTCAATTTGAATTAATTAATTAAGTCCGACATAGTCAACACACACACA	2119

Db	17757	AAACAATTGTTCAAAATTGTAAATACTATTAAAGTCGACACTGTCAAAACAACAACA	17698
Qy	2120	GCACACGAAAAACAACAAAAAAATATGAAAAACACACAGAAAGAACCTTCAATTG	2179
Db	17697	GCAACAGCAAAAAACAACAAAAAAATGTGAAAAACAAGCAAGAACCAATTCATTCAG	17638
Qy	2180	ATCATTTAGCAATCGATGTTAAATTAATTAATTACTAAAGTCACCTTAATGCTTA	2239
Db	17637	ATCATTTAGCAATCGATGTTAAATTAATTAATTAACTAAAGTCACCTTAATGCTTA	17578
Qy	2240	CAAAATGAGCATGCTAATCCCTTACACACACACACACACACACACTCGAAAGTTTA	2299
Db	17577	CAAAATGAGCATGCTAATCCCTTACACACACACACACACACACACTCGAAAGTTTA	17518
Qy	2300	CTAATTTAATTTATTTATGTTAGGGACGAGGCTTATTAATTGCTCAATTGAGGAAC	2359
Db	17517	CTAATTTAATTTATTTATGTTAGGGACGAGGCTTATTAATTGCTCAATTGAGGAAC	17458
Qy	2360	TATTTATTTATTTATTTATTTAATTAATTTAGTGAATTCACACAAACAGCAGAAAAA	2419
Db	17457	TATTTATTTATTTATTTATTTAATTTAGTGAATTCACACAAACAGCAGAAAAA	17398
Qy	2420	CAACACAACAACAAGAGAGAGAGAAAACAAACCAATTCACCTGTAATAAATATCCATTGA	2479
Db	17337	CAACACAACAACAAGAGAGAGAGAAAACAAACCAATTCACCTGTAATAAATATCCATTGA	17338
Qy	2480	AAAAATACACAGAAAAAGCAAGAAATTAATAAAATCAAAACATTTCAAGATACAAAGT	2539
Db	17337	AAAAATACACAGAAAAAGCAAGAAATTAATAAAATCAAAACATTTCAAGATACAAAGT	17278
Qy	2540	AATTAACAAAATTAACAGCAA 2558	
Db	17277	AATTAACAAAATTAACAGCAA 17259	

RESULT 4	AC010564	LOCUS	DEFINITION
	AC010564	196594 bp	DNA linear INV 01-JUN-2002
	<i>Drosophila melanogaster</i> 3L BAC RP98-2701 (Roswell Park Cancer Institute <i>Drosophila</i> BAC Library) complete sequence.		

ACCESSION	AC010564	GI:21306584
VERSION	AC010564.7	
KEYWORDS	HTG.	
SOURCE	<i>Drosophila melanogaster</i> (fruit fly)	
ORGANISM	<i>Drosophila melanogaster</i>	

REFERENCE	1 (bases 1 to 196594)
AUTHORS	Muzny, D., Scherer, S., Adams, M.D., Holt, R.A., Evans, C.A.,

Dugan-Rocha, S. D., Sodergren, E. S., Hodgson, A. H., Chen, R. C., Ayale, M., Scott, G. S., Worley, K. W., Anamathides, P. G., Brandon, R. C., Rogers, Y., An, H., Baldwin, D., Beeson, K. Y., Brown, M., Butthy, C., Bussem, D. A., Center, A., Chen, G., Chen, Z., Clerc-Blankenburg, K., Davisport, L. B., Dietz, S. M., Ding, Y., Dodson, K., Doup, L. E., Draper, H., Emery-Cohen, A., Ferriere, S., Gary, N. D. S., Houck, J., Hostin, D., Howland, T. J., Hume, J., Ibegwam, C., Jallali, M., Kovar, C., Liu, W., Metcalf, B., McIntosh, T. C., Morgan, M., Noy, M., Murphy, B., Nelson, K. A., Ndassia, Y., Nguyen, N., Perez, L., Piltman, G. S., Furi, V., Schaefer, F., Shen, H., Strong, R., Tector, C., Wang, O., Williams, S. M., Xiang, J., Zaveri, J. S., Zhou, J., Zorrilla, S., Smith, H. O., Wheeler, D., Weinstock, G., Gibbs, R. and Venter, J. C.

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 196594)
Worley, K.C., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Albrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbara, J.,
Benton, J., Blincoe, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burich, P., Burrell, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carlier, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, T., Chowhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,

TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	3 (basee 1 to 196594)
AUTHORS	Morley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (16-SEP-1999)
REFERENCE	Human Genome Sequencing Center, Department
AUTHORS	of Molecular and Human Genetics, Baylor College of Medicine, One
TITLE	Baylor Plaza, Houston, TX 77030, USA
JOURNAL	4 (basee 1 to 196594)
REFERENCE	
AUTHORS	Morley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Bentley, J., Dingy, N., Brown, E., Brown, M., Bryant, N. P., Bunhay, C.,
Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N. P., Bunhay, C.,
Burck, P., Burkett, C., Burrell, K. L., Byrd, N. C., Carron, T. F.,
Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C. D., Cox, C.,
Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C.,
Davy-Carroll, L., Dederich, D. A., Delaney, K. R., Delgado, O.,
Dean, A. L., Ding, Y., Dinh, H. H., Douthwaite, K. J., Draper, H.,
Dugan-Rocha, S., Durbin, K. J., Sarnath, C., Edgar, D., Edwards, C. C.,
Elhaj, C., Escoto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrill, J. H., Guevara, W., Gunaratne, P., Hale, S.
Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C.,
Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hune, J.,
Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Joivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratoch, J., Kuchel, A., Landry, N., Leal, B., Lewis, L. C.,
Lewis, L., Li, J., Li, Z., Kitchashi, O., Ilean, C., Liu, J., Liu, M.,
Loulleged, H., Lozado, R. J., Lu, X., Lucier, A., Lucier R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Matlindale, A.,
Martinez, E., Massey, E., Mawhinney, E., McLeod, M. P., Meador, M.,
Mei, G., Mercker, M., Miner, G., Minier, Z., Mitchell, T., Monobacht, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokekwo, S.,
Ogulu, M., Okunolu, G., Orazungu, N., Oviedo, R., Pace, A., Payton, B.,
Pugh, J., Perez, L., Peters, L., Plickens, R., Pirmus, E., Pu, L. J.,
Quilas, M., Ren, Y., Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M.,
Rutiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoochari, N.,
Stison, I., Sodergren, E., Sonalle, T., Sparks, A., Stanley, H.,

[illegible]

Accession	Gene	Protein	Accession	Gene	Protein
QY 2420	AAACAAACAACGAGAGAGAGAAA	CAAAACCAATTCACGTAAATATCCAAATTGA	2479		
Db 112682	CAACCAACAACAAGAGAGAGAAA	CAAAACCAATTCACGTAAATATCCAAATTGA	112741		
QY 2480	AAATATACACAGAAAAGCCAAAGAAAT	ATCAAAATCAATTCAGAGATACCAACAGT	2539		
Db 112742	AAAATATACACAGAAAAGCCAAAGAAAT	ATCAAAATCAATTCAGAGATACCAACAGT	112801		
QY 2540	AATAACAAAATATACAAAGCAA	2558			
Db 112802	AATAACAAAATATACAAAGCAA	112820			

RESULT 5
AE003471

LOCUS AE003471 302131 bp DNA linear INV 14-FEB-2003
DEFINITION Drosophila melanogaster chromosome 3L, section 5 of 83 of the
complete sequence.
ACCESSION AE003471 AE002584 AE014296
VERSION AE003471.3 GI:23092741

ORGANIZATIONAL

REFERENCE
AUTHORS

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 302131)

Adams, M.D., Celisner, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D.,

TITLE
JOURNAL

Man, K.H., Doyle, C., Baxter, E.G., Heltz, G., Nelson, C.R., Gabor, G.L.,
Abill, J.F., Agbeyanti, A., An, H.J., Andrews-Pfannkoch, C., Baldwin,
B., Ballew, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M.,
Beeson, K.Y., Benos, P.V., Bernier, B.P., Bhattacharya, D., Bolintsov, S.,
Borkova, D., Borchan, M.R., Bouck, J., Brockstein, P., Brothier, P.,
Burtis, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A.,
Chandra, A., Cherry, J.M., Cawley, S., Dahlke, C., Daveport, L.B.,
Davies, P., de Pablos, B., Delcher, A., Deng, Z., May, A.D., De, I.,
Diez, S.M., Dodson, K., Dou, L.E., Downes, M., Dugan-Rocha, S.,
Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C.,
Ferreira, S., Fleischmann, W., Foster, C., Gabrielian, A.E., Garg, N.S.,
Guan, P., Harris, M., Glasser, K., Glodde, A., Gong, F., Gorrell, J.H., Gu, Z.,
Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J.,
Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J.,
Wei, M.H., Iadegam, C., Jaisli, M., Kalush, F., Karpen, G.H., Ke, Z.,
Kensimon, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Krat, C.,
Krivitz, S., Kulp, D., Lai, Z., Lasro, P., Lei, Y., Levitsky, A.A.,
Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C.,
McLeod, M.P., McPherson, D., Meekunov, G., Mishina, N.V., Murphy, C.,
Morris, J.P., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L.,
Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K.,
Nusskern, D.R., Pacleb, J.M., Palazolo, M., Pictman, G.A., Pison, P.,
Pollard, J., Furi, V., Reese, M.G., Reinert, K., Remington, K.,
Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I.,
Simpson, M., Skupski, M.P., Smith, T., Spiller, E., Spradling, A.C.,
Steptleton, M., Strong, R., Sun, E., Svrtkars, R., Tector, C., Turner, R.,
Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A.,
Weinsteck, G.M., Weissensbach, J., Williams, S.M., Woodaet,
Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yen, R. P.,
Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H.,
Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O.,
Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.

The genome sequence of *Drosophila melanogaster*
Journal 287 (5461), 2185-2195 (2000)

REFERENCES

Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y.,
Banzon, J., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A.,

CARLSON,J.W., CENTER,A., CHAMPE,M., DAVENPORT,L.B., DIETZ,S.M.,
DODSON,K., DORSETT,V., DOUP,L.B., DOYLE,C., DRENEK,D., FARFAN,D.,
FERREIRA,S., FRISSE,E., GALLIE,R.F., GARG,N.S., GEORGE,R.A.,
GONZALEZ,M., HOUCH,J., HOSKINS,R.A., HOSTIN,D., HOWLAND,T.J.,
IBEGWAM,C., JALALI,M., KRUSE,D., LI,P., MATTEI,B., MOSHREI,A.,
MCINOSH,T.C., MOY,M., MURPHY,B., NELSON,C., NELSON,K.A., NUNOO,J.,
PACLEB,J., PARAGAS,V., PARK,S., PATEL,S., PFEIFFER,B.,
PHOUNENAVONG,S., PITTMAN,G.S., PUTI,V., RICHARDS,S., SCHEELER,F.,
STAPLETON,M., STRONG,R., SVILSKAS,R., TECTOR,C., TYLER,D.,
WILLIAMS,S.M., ZAVARI,J.S., SMITH,H.O., VENTER,J.C. and RUBIN,G.M.
Sequencing of Drosophila melanogaster genome
Unpublished
3 (bases 1 to 302131)
REFERENCE
AUTHORS
Mista,S., Crosby,M.A., Matthews,B.B., Bayraktaroglu,L.,
Campbell,K., Hradecky,P., Huang,Y., Kaminker,J.S., Prochnik,S.E.,
Smith,C.D., Tupy,J.L., Bergman,C.M., Berman,B.P., Carlson,J.W.,
Celniker,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Frisze,E., de
Grey,A.D.N.J., Harris,N.U., Krommiller,B., Marshall,B.,
Millburn,G.H., Richter,J., Russo,S., Seale,S.M.J., Smith,E.,
Shu,S., Smutnajak,F., Whitfield,E.J., Ashburner,M., Gelbart,W.M.,
Rubin,G.M., Mungall,C.J. and Lewis,S.B.
Annotation of Drosophila melanogaster genome
Unpublished
4 (bases 1 to 302131)
REFERENCE
AUTHORS
Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
5 (bases 1 to 302131)
REFERENCE
AUTHORS
CONSRM
TITLE
JOURNAL
FlyBase
Direct Submission
Submitted (06-SEP-2002) University of California Berkeley, 539 Life
Sciences Addition, Berkeley, CA 94720, USA
6 (bases 1 to 302131)
REFERENCE
AUTHORS
CONSRM
TITLE
JOURNAL
FlyBase
Direct Submission
Submitted (23-JAN-2003) University of California Berkeley, 539 Life
Sciences Addition, Berkeley, CA 94720, USA
On Sep 17, 2002 this sequence version replaced gi:10727219.
Location/Qualifiers
1. 302131
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="3L"
/note="genotype: y[1]; cn[1] bw[1] sp[1]; Rh6[1]"
277. .2606
/locus_tag="CG1211"
/note="last curated on Thu May 16 11:24:41 PDT 2002"
/map="61P4-61P4"
/db_xref="FLYBASE:FBgn0035191"
join(277. .596,650. .1022,1078. .1302,1405. .1847,1913. .2606)
/locus_tag="CG1211"
/product="CG1211-RA"
/db_xref="FLYBASE:FBgn0035191"
join(277. .596,650. .1022,1078. .1302,1405. .1847,1913. .2606)
/locus_tag="CG1211"
/note="CG1211 gene product"
/codon_start=1
/product="CG1211-PA"
/protein_id="AA047455.1"
/db_xref="GI:7292042"
/translation="MSGSSPNSIRPLPGIQKVAIBELNEVPSRVSDIALAKEWLQK
OHLACLEDOFLPLGSKFSLKAKOKIDRFYSLQAVIPFVNDRLVNDVAYLE
IRLGVILRLIDEDDTGPAYITIRAGSYDINKERODIIRVSGMGTIMLMDNNA
VSGYLIMMSGCTGANNLRLQOLSKSAYADAMPTRQGIHPIINPKAPETGFK
SLGMPFGKIKERKVSVDPEALFERVGHVLPPEEGSKGTWKDITDWEALCSYR
SYFEDQHGAKHDKREASVLPNDSEHGLTLGLGYPVWMSALSRLIDFGSKR
TWWKQHSFIDQEBRNKMPAIRPLSPLEOKAIINLNEVPLKDLDAALRWKQ
OHLKARTDQFLVNFRLRGCKFSLERTKSKIDRFYLRKYPFYLGHNVDKALEI

FLGIVTILPRELNDNGRLALLIRNACYPDSKYTLQEVNRAAGLMQOIMLDEDDVAIV
NGLISILDSNVTGTHFLQMSPSFAKKTIVFOEALPIRPOGIHPIINPENGDTIFNM
IKPMKQOQSGNSPNCOTKSMKTYLILHQLVYHSGKMWALNQIIPKOYLPIVYG
GENGSIPELQOQMEORILALYRNRYWEKNVGTDESLRNGQPDPSFLGLQSGFRQLN
VD"
3102. .9398
/locus_tag="CG9194"
/note="last curated on Thu May 16 11:26:33 PDT 2002"
/map="61P4-61P4"
/db_xref="FLYBASE:FBgn0035192"
join(3102. .3286,3891. .4150,4210. .4456,7001. .7127,
7192. .8594,8650. .9398)
/locus_tag="CG9194"
/product="CG9194-RA"
/db_xref="FLYBASE:FBgn0035192"
join(3944. .4150,4210. .4456,7001. .7127,7192. .8594,
8650. .8855)
/locus_tag="CG9194"
/note="CG9194 gene product"
/codon_start=1
/product="CG9194-PA"
/protein_id="AA047456.1"
/db_xref="GI:7292043"
/translation="MSGRAQSLPAMHLEPAKPORGCVAALCPSMKVLTCIVSHVL
VLLVYCVGAYLFOHLERPHLEKEDIOILRVNLEINILDLSDAVLESDHMA
NYSKLAPNEKOILPAIKADGMDGDEDRKSGMTPAGSLFYISITITIGVHSEPT
DMGKTTTIFYAVIGVPLMLICLSNIGDMATYSFRILYRICCYVTRTRAKRNASR
QKMSRQSHAKQPPPSFRSMKTKQKSGNSGLSPSGHATSDPLRTMGQYDRE
FGRSSGGGRNRQOQOHLHDHPRQRTITVDGYETQLNNSNYSSRQGRNDIRD
RHTVERERYSRHLNAGSIEDGDMQPAKRAASVRSVSPINQSSKASRELIHLS
AAGRSKASVDPRIHSHYEDVEDVYKPTIIPNRVALLDPCGRROAAPSQMPR
SAHOROKROKREERSPOPPQSSVROOHDRAASLGROSSRYGNHLIPDYDAPPG
BDHRDRRGSGKRYEDYVESFDEGSLXGDNNYDYPERRHSRREKRNRERA
EKLPPSPRIMSPMGFPVQRQIKRPSYTDIDDDSMYGDYGDLPDRVPYIWL
VFLVSYLIGAVLPAVWMSFLDSAYFCFTITLTITIGGDFPKAGVDSBSOIAV
CSLYLFGIALAMSPNLVQBEFFIANVEXARLIGILKDDDDDEQDED"
/locus_tag="CG9192"
/note="CG9192"
/locus_tag="CG9192"
/note="last curated on Wed May 22 08:07:26 PDT 2002"
/map="61P4-61P4"
/db_xref="FLYBASE:FBgn0035193"
/locus_tag="CG9192"
/product="CG9192-RA"
/db_xref="FLYBASE:FBgn0035193"
/locus_tag="CG9192"
/note="CG9192 gene product"
/codon_start=1
/product="CG9192-PA"
/protein_id="AA047457.1"
/db_xref="GI:7292044"
/db_xref="FLYBASE:FBgn0035193"
/translation="MPSYCLITLCLLSVAKASDGDGRDQVMSYNRRESRFPDIN
RSPPIRNSYDEKLAREPPTTRPPEPSPNDIEDVADVPKRSPTPVNTRKSPYNG
GQTNAQGGGLAGGYNNGVNLVGPBGPTGIIIGRPDYTPPYPGYGGSGAONISGY
PEGYGVQGLGAGCFPGASVGGGFTNNPNSNLGNLPGNMQTYGISTPGSDPFGSG
FAGAGAQPGLGQPANQPGPOYTBXYGLAGLGLQOLGVNGGLGYGGSFPLPG
GFGYBKSPAVADGSKAKSVASPRKXNDKSKKV"
complement(8941. .13312)
/gene="klp61F"
/locus_tag="CG9191"
/note="last curated on Wed May 22 07:37:48 PDT 2002"
/map="61P4-61P4"
/db_xref="FLYBASE:FBgn0004378"
complement(join(8941. .9364,9449. .9700,9765. .10975,
11041. .12373,12567. .12867,12934. .13312))
/gene="klp61F"
/locus_tag="CG9191"

Query Match 67.4%; Score 1723.4; DB 3; Length 302131;
Best Local Similarity 93.3%; Pred. No. 0;

Matches 1866; Conservative 0; Mismatches 1; Indels 132; Gaps 1;

```
Qy 692 TTGAGATTGCACTTCGCTTACGACCGCTACCAATATGCGCGCAAGATTTGGGGCTAC 751
Db 23358 TTTAGATTGCACTTCGCTTACGACCGCTACCAATATGCGCGCAAGATTTGGGGCTAC 233917
Qy 752 CCGTTCGATTCGCTGCGATTCGCTGCTGCTATCGCGCGGACCGCGCTGCGAGGTGT 811
Db 233918 CCGTTCGATTCGCTGCGATTCGCTGCTGCTATCGCGCGGACCGCGCTGCGAGGTGT 233977
Qy 812 GCGGCTTCCTTACGATGTTCTGCGACGCGCAATGCTGCTACCTGCGCTTCAATATG 871
Db 233978 GCGGCTTCCTTACGATGTTCTGCGACGCGCAATGCTGCTACCTGCGCTTCAATATG 234037
Qy 872 TCATTCAGCTGTTTGGGATTCCTCGAGAGGATGCAAGGACGCGGACGAGATGCGGCG 931
Db 234038 TCATTCAGCTGTTTGGGATTCCTCGAGAGGATGCAAGGACGCGGACGAGATGCGGCG 234097
Qy 932 TGATCTACATGCGCGGCTTTTGGCCGATCCCTGGGACCAAGTGTGCTGCACTCGAGG 991
Db 234098 TGATCTACATGCGCGGCTTTTGGCCGATCCCTGGGACCAAGTGTGCTGCACTCGAGG 234157
Qy 992 TCTTCCTGCTGGGCGCGAGCGGTGCGCTCTATGCTGCTGTTGGCCGCACTGCGCAACA 1051
Db 234158 TCTTCCTGCTGGGCGCGAGCGGTGCGCTCTATGCTGCTGTTGGCCGCACTGCGCAACA 234217
Qy 1052 TTTACACTGATGAGGCGACATGAGAGCGCATCCAGCAATCGATTCGCTGCTGCTACT 1111
Db 234218 TTTACACTGATGAGGCGACATGAGAGCGCATCCAGCAATCGATTCGCTGCTGCTACT 234277
Qy 1112 TT----- 1113
Db 234278 TTGGTGAAGTTTCCCAACCTTATGACTAGATCTATAGAACTTATCTGACTTAA 234337
Qy 1114 ----- 1113
Db 234338 ACAAGATTTGTCACTTTTCTTCTATGATTTAACTTAACTTGAATCCCTTCTG 234397
Qy 1114 -----GTTCTCGTGGATCTGGGCTATGCTCTCTACACCAATCTGCTGATG 1159
Db 234398 TTCTTCATTTGACGATCTCTGCGATCGGCTATGCTCTCTACACCAATCTGCTGATG 234457
Qy 1160 GAAGCGGCTTCGCGCAAGGTCCTCCAGGTCGTATCATTTGCCACTGACGCGAGCCCTG 1219
Db 234458 GAAGCGGCTTCGCGCAAGGTCCTCCAGGTCGTATCATTTGCCACTGACGCGAGCCCTG 234517
Qy 1220 CAGGACTAACGATCGCTTCTGCTGCTAAAGAACTTCGCTCACGAGAGTACGAGCAGC 1279
Db 234518 CAGGACTAACGATCGCTTCTGCTGCTAAAGAACTTCGCTCACGAGAGTACGAGCAGC 234577
Qy 1280 TCATCTGCTGCTAGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339
Db 234578 TCATCTGCTGCTAGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 234637
Qy 1340 ACCGATCAACACGCTGACCGCTCAGCTGATGAGAGAGAGGAGTATTAACGACG 1399
Db 234638 ACCGATCAACACGCTGACCGCTCAGCTGATGAGAGAGAGGAGTATTAACGACG 234697
Qy 1400 ATCTGTTGCAAGACTGCGAGGTGCTCTAAGTTTGAAGTTTGGAGTCTGCAAGTCTGCG 1459
Db 234698 ATCTGTTGCAAGACTGCGAGGTGCTCTAAGTTTGAAGTTTGGAGTCTGCAAGTCTGCG 234757
Qy 1460 AGGATTCGGAATCTGCTTGAAGCTTCAAGAGAGATCGAGAGACAGAGAGTTGGTGAAGA 1519
Db 234758 AGGATTCGGAATCTGCTTGAAGCTTCAAGAGAGATCGAGAGACAGAGAGTTGGTGAAGA 234817
Qy 1520 GAAAAGTTCACTCAAGATTTAGTTCAAAAATAATGATTTGTTTGGCTTTT 1579
Db 234818 GAAAAGTTCACTCAAGATTTAGTTCAAAAATAATGATTTGTTTGGCTTTT 234877
Qy 1580 TCGTTACCATTTATCTGTTATCTGTTATCTGTTACGTTTGAAGTTTAAAGTTTCCGAAACAT 1639
Db 234878 TCGTTACCATTTATCTGTTATCTGTTATCTGTTACGTTTGAAGTTTAAAGTTTCCGAAACAT 234937
```

```
Qy 1640 AGTACCAAACTCAAAAAAAAAAAAAAAAACAAAATCAAGAGAAATACCTGACAAAAAGAAA 1699
Db 234938 AGTACCAAACTCAAAAAAAAAAAAAAAAACAAAATCAAGAGAAATACCTGACAAAAAGAAA 234997
Qy 1700 GTGCGAGAGTGAAGAGAACTATTAACCGAAGCCGAAAGCTGTAAACAAATGTTGTGATG 1759
Db 234998 GTGCGAGAGTGAAGAGAACTATTAACCGAAGCCGAAAGCTGTAAACAAATGTTGTGATG 235057
Qy 1760 AACCAAGAAGTGAATTTATTTGCGGTAAAAAACAAGTAAAAATCAAGAGGAAATCA 1819
Db 235058 AACCAAGAAGTGAATTTATTTGCGGTAAAAAACAAGTAAAAATCAAGAGGAAATCA 235117
Qy 1820 AGAGAGAAACAGAACTAAATGCTCTCCGCTATGATTTAAATGATTTATCAATGTTTCA 1879
Db 235118 AGAGAGAAACAGAACTAAATGCTCTCCGCTATGATTTAAATGATTTATCAATGTTTCA 235177
Qy 1880 ATTAATGTTTGTCTGTTTCTTAAATTAATGATTTATTTAGTGGCCGCAATTAACGATGA 1939
Db 235178 ATTAATGTTTGTCTGTTTCTTAAATTAATGATTTATTTAGTGGCCGCAATTAACGATGA 235237
Qy 1940 ATGGAATGGAAGATGAGAACTGATCAATGTTATGATTAATGATTAAGCATATATG 1999
Db 235238 ATGGAATGGAAGATGAGAACTGATCAATGTTATGATTAATGATTAAGCATATATG 235297
Qy 2000 CTCGAATTTGAGATTTATGTTATTAATTTATTAATTTAAGTAACTAGCCCTCT 2059
Db 235298 CTCGAATTTGAGATTTATGTTATTAATTTATTAATTTAAGTAACTAGCCCTCT 235357
Qy 2060 AACAAATGTTTCAAAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2119
Db 235358 AACAAATGTTTCAAAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 235417
Qy 2120 GCAACGCAAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 2179
Db 235418 GCAACGCAAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 235477
Qy 2180 ATCAATTAAGCAAAATGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2239
Db 235478 ATCAATTAAGCAAAATGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 235537
Qy 2240 GAAATTCAGACATGTAATCCCTTACACACACACACACACACACACACACACACAC 2299
Db 235538 GAAATTCAGACATGTAATCCCTTACACACACACACACACACACACACACACACAC 235597
Qy 2300 CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2359
Db 235598 CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 23657
Qy 2360 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2419
Db 235658 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 235717
Qy 2420 CAACCAACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2479
Db 235718 CAACCAACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 235777
Qy 2480 AAAATACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2539
Db 235778 AAAATACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 235837
Qy 2540 AATACAAATAATACAGCAA 2558
Db 235838 AATACAAATAATACAGCAA 235856
```

RESULT 6
AC008252/c 4985 bp DNA linear HTG 02-AUG-1999
LOCUS
DEFINITION
Drosophila melanogaster chromosome 3 clone DS05739 (D460) map
62A3-62A5 strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***, 40
unordered pieces.
ACCESSION AC008252

VERSION	KEYWORDS
AC008252.1	GI:5656704
SOURCE	ORGANISM
HTG: HTGS_PHASE1	Drosophila melanogaster (fruit fly)
REFERENCE	AUTHORS
1 (bases 1 to 49858)	Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazey,R.G., Burenhoff,C., Chambe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotar,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Peclab,J.M., Park,S., Pfeiffer,B., Poon,L., Seguire,A., Sethi,H., Snir,E., Svirkas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.
TITLE	JOURNAL
Sequencing of Drosophila melanogaster	Unpublished
REFERENCE	AUTHORS
2 (bases 1 to 49858)	Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazey,R.G., Burenhoff,C., Chambe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotar,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Peclab,J.M., Park,S., Pfeiffer,B., Poon,L., Seguire,A., Sethi,H., Snir,E., Svirkas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.
TITLE	JOURNAL
Direct Submission	Submitted (30-JUL-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-111, Berkeley, CA 94720, USA
COMMENT	For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (http://www.fruitfly.org/sequence/) or send email to bdg@fruitfly.berkeley.edu . All contigs in this submission meet the following cutoffs: length >= 200 bases. Pl library location: 60-75.
* NOTE: This is a 'working draft' sequence. It currently	
* consists of 40 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
* 1	821: contig of 821 bp in length
* 822	901: gap of unknown length
* 902	1211: contig of 310 bp in length
* 1212	1291: gap of unknown length
* 1292	2028: contig of 737 bp in length
* 2029	2108: gap of unknown length
* 2109	2765: contig of 657 bp in length
* 2766	2845: gap of unknown length
* 2846	3638: contig of 793 bp in length
* 3639	3718: gap of unknown length
* 3719	4386: contig of 668 bp in length
* 4387	4466: gap of unknown length
* 4467	5393: contig of 927 bp in length
* 5394	5473: gap of unknown length
* 5474	6220: contig of 747 bp in length
* 6221	6300: gap of unknown length
* 6301	7269: contig of 969 bp in length
* 7270	7349: gap of unknown length
* 7350	8466: contig of 1117 bp in length
* 8467	8546: gap of unknown length
* 8547	9543: contig of 997 bp in length
* 9544	9623: gap of unknown length
* 9624	10675: contig of 1052 bp in length
* 10676	10755: gap of unknown length
* 10756	11702: contig of 947 bp in length
* 11703	11782: gap of unknown length
* 11783	12934: contig of 1152 bp in length

[illegible]

QY 1026 CTTGTGGCCGCACTCTGSCCAATTTACATGAACTATGCGACATGAAAGCCGATC 1085
 |||||
 DB 1015 CTTCTGCGCCGCACTTTGGCCAACTTACGTGAATATGCGACATGAAGTCCGATC 1074
 |||||
 QY 1086 CAGGCACTCGGATTCCTTTGTCATCTTTCTCTGGAATCTGGGCTATGCTCTTACAC 1145
 |||||
 DB 1075 CAGGCACTCGGCTCTGTGTGTCATTTTGTCTCTGGAATCTGGGCTATGCTCTTACAC 1134
 |||||
 QY 1146 CCAATCTTGTGATG-----AAGCGCTTCGCGAAGGATCCCGAGGTGTCAT 1196
 |||||
 DB 1135 GCAATCTTGTGATGCGCTGCGCTGCGACGCCGCTTTGCGAAGGCGCCCGAGTCTGATAT 1194
 |||||
 QY 1197 TGGCCACCTTACCGGAGCCCTGCGACGACTTACGATGAGCTTTCTGTGCTTAAAGACT 1256
 |||||
 DB 1195 AGGCACTTACCGGCGCTCTGCGCGGCTGACCATGTGCTTTCTGTGCTTCAAGATTT 1254
 |||||
 QY 1257 CGGTCACTGAGATGAGAGCTCATCTGTGCTTACGCTTGGGCGCTCTTCTGCTT 1316
 |||||
 DB 1255 TGGACATCTGGAATACCAACGATGATCTGTGCTGCGCTGCGCTGCTTATTTGCTT 1314
 |||||
 QY 1317 CACGCTTCTGCGCATGCTTTTCAACCTGATCAACAGGCTGACGCTGAGTGAAGAG 1376
 |||||
 DB 1315 CACGCTTCTGCGCATGCTTTTCAACCTGATCAACAGGCTGACGCTGAGTGAAGAG 1374
 |||||
 QY 1377 GCAAGGTGAGTGAATTAACCGACATCTGTGCAAGACTTGGAGTGTCTTAACTTTGAGG 1436
 |||||
 DB 1375 AGAGGGGAGGTGATGACACATTTGCTGCAAGCTTGGGCGTGTCTTAAAGCAACAG 1434
 |||||
 QY 1437 TTGCGAGTCTGACGA---TGCTCGAGGAGATTCGGAATCTGCTTGAAGCTTACGAGAG 1492
 |||||
 DB 1435 GCAATATATACAGAAACCAACAAACAAATATTTCAAGAGAGAGAAACATAGACAA 1494
 |||||
 QY 1493 ATGAGAGACAGAGAGTGTGAGAAAGAAAGTTCACCTCAAGATTTAGTCAAAACTA 1552
 |||||
 DB 1495 AAAAAAACAAGAAAGTGAAGAAAGAAAGTGAAGAGAAAGTTAAATCGAAACCTG 1554
 |||||
 QY 1553 ATT-----CGATATTCGTTTGGCTTTTCTGCTTCTTGAATCTGCTTATCTGTTAGCTTA 1605
 |||||
 DB 1555 ATTTTATGTTACAAACTTAATTTGAAATTCGTTGCTTAACTTTTCTGCTTCCGTT 1614
 |||||
 QY 1606 CCGTTTGAAGTAAAGCTTTCACTTCCGAAACATAGTACCAAACTCAAAAAA 1665
 |||||
 DB 1615 CCGTTTCAATTCGTTCCGTTGCGTTCCGTTTAACTGTAAGCGGCTTTCATTTCCACAT 1674
 |||||
 QY 1666 ACAATATCAAGAAATTAATCACTGACCAAAAGAAAGTCCGAGAGTGAAGAACTATTAAC 1725
 |||||
 DB 1675 TCAAGCAAGAAAGAACCAACCAACCAACCAACCAAGAAAAAGAACTACTCA 1734
 |||||
 QY 1726 CGAA---GCCGAACGTGTAAACAATGTTGTGATAGAACCAAGACTGAATTTATTTG 1782
 |||||
 DB 1735 AGAAGCGGCGGAAAGCTTAAACAAATGTTGTGATAGAACCAAGACTGAATTTATTTG 1794
 |||||
 QY 1783 CGT 1785
 |||||
 DB 1795 CGT 1797
 |||||

RESULT 8
 AY118997 1816 bp mRNA linear INV 16-JUN-2002
 LOCUS Drosophila melanogaster LP02893 full insert cDNA.
 DEFINITION AY118997
 ACCESSION AY118997.1 GI:21430357
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Phyllophoridae; Drosophilidae; Drosophila.
 1 (bases 1 to 1816)
 Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,
 Champe, R., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Fries, E.,
 George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G.,

Miranda, A., Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
 Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Rubin, G.M.
 and Celisier, S.
 Direct Submission
 Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,
 Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, USA
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This clone was sequenced as part of a high-throughput process to
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
 Science 2000). The sequence has been subjected to integrity checks
 for sequence accuracy, presence of a polyA tail and contiguity
 within 100 kb in the genome. Thus we believe the sequence to
 reflect accurately this particular cDNA clone. However, there are
 artifacts associated with the generation of cDNA clones that may
 have not been detected in our initial analyses such as internal
 priming, priming from contaminating genomic DNA, retained introns
 due to reverse transcription of unspliced precursor RNAs, and
 reverse transcriptase errors that result in single base changes.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our web site
 (<http://fruitfly.berkeley.edu>) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES

source

gene

CDS

1..1816
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /strain="Y; cn bw sp"
 /db_xref="taxon:7227"
 1..1816
 /gene="ru"
 /db_xref="FLYBASE:FBgn0003295"
 320..1345
 /gene="ru"
 /note="Longest ORF"
 /codon_start=1
 /product="LP02893p"
 /protein_id="AAM50857.1"
 /db_xref="GI:21430358"
 /db_xref="FLYBASE:FBgn0003295"
 /translation="MILSGAPSDGRRPRLNRLIGTVERGEMPSVTLQMPAPSSK
 SGLVGVCCDQIMAVOPNORASGAATTSNSDPMKNNRKHGGSAAPKMTPEPTI
 LATTLELVFWGADPPEDSLVTRPPDRLOMPFLSLALHSHSLHGVTVTL
 FVSPLELVHGSRTGVITVAGVLAGSLGTVSDSEFLVAGSGVYALIAQALSL
 NFGQMRHGVITQMAVILFVFCIDYGLVSRRLAMHQLQTRPSVSIYAHNTGLAGISV
 GILLRLQDGGRLPRPLRLALGVWCIFSAFGLIARLVTVTAQLAEBEGVIRHIL
 NMDLGMG"
 BASE COUNT 448 a 481 c 452 g 435 t
 ORIGIN

Query Match 9.8%; Score 251.8; DB 3; Length 1816;
 Best local Similarity 62.3%; Pred. No. 1.9e-40;
 Matches 413; Conservative 0; Mismatches 247; Indels 3; Gaps 1;

QY 763 CCGTGGATTGGTGTCTTATGCGCCGACCGGCGTGTGAGGTGGCGCTTTCTT 822
 |||||
 DB 677 CCAAGAGATTCCTTCTGTATATGACCCGACACGCGCTGACGCTGAGATTCCTT 736
 |||||
 QY 823 AGCTACATTCCTGTGAGCGCAACTGCTTCACTGCGGCTTCAATATGCTATCCAGCTG 882
 |||||
 DB 737 AGCTATGCGCTGTGACCGCAAGTGTGCTTCACTGCGGCTTCAATATGCTATCCAGCTG 796
 |||||
 QY 883 TTTTGGCATTTCCCTGAGGATGATGACGCGACGCGCAGGATCGGCTGATCTTACATG 942
 |||||
 DB 797 CTTTGGATGATGACATGAGCTGTGATGATGCTGAGAACGGAATCTTATATAG 856
 |||||
 QY 943 GCGGCGTTTGGCGGATCCTGCGCACAGTGTCTGACTCGGAGTCTTCTGCTG 1002
 |||||
 DB 857 GCGGAGTTTGGCTGCTCTTGGGAACTTGGTGTGACTCGGAAATTTTCTGCTG 916
 |||||

QY	1003	GGCGCACGGGCGCTGATAGCCCTGTTGGCCGCAATCTGGCCCAATTAACCTGAAC	1062
Db	917	GGAGCGAATGAGGAGTTTACGCCCTTGCTGGCCGCCCACTGGGAGCTTTTGCTCAAC	976
QY	1063	TATGCGCAATGAAGAGCGCATCCAGCACTCGGATCGTTTGATCTTGTCTCTGC	11222
Db	977	TTTGGTCAAAAGAGACACGGGGTAATCAGCTTAATAGCTGTGATTTTGTGTTTCTGC	1036
QY	1123	GATCGGGGCTATGCTCTCTACACCCAACTACTCATGGAAGCGCCCTTGCCAAAGGCTCC	1182
Db	1037	GATTTGGGATATGCCCTCTATTCAAGAAACTGGCAATGATCAAGCTGCAACACAGACC	1096
QY	1183	CAGGTGTGTAACATTTGCCACCACTGACCGGAAGCCCTGGAGGACTAAACGATGGCTTTCTG	1242
Db	1097	TCTGTCTCTATATGCTCACTGACCGGTGCTGGCCGGGATTAAGTGGGTCGTG	1156
QY	1243	GTGCTAAAGAACTTGGTGATCCGGAAGTACAGAGAGCTCATCTGTGGCTACGTTGGGC	1302
Db	1157	CTTCTCCGGCAACTGGATGTGGGCTGGACCAAGCCCACTCCGTGGTTGGCCCTTGGC	1216
QY	1303	GTCTACTGTGACCTTCAACGCTCTTGCSCATCGTTTCACTCGATCAACACGGTAGCCGT	1362
Db	1217	GTAATGTGCATTTATGAGCGCCTTTGGAAATCCGCTTCAACTTGTGTAAATACGGTCAACGT	1276
QY	1363	CAGCTGATG--GAGAGACAGGGTAGGTGATTAACCAAGCATCTGTTGACAGACTTGGA	1419
Db	1277	CAACTTCTGGCCGAGAGAGAGGGCCAGGTGATCAGGCGAGCATCTGATGAATGAACTCCGA	1336
QY	1420	GTG 1422	
Db	1337	ATG 1339	

RESULT 9	
AF318284	1822 bp mRNA linear INV 11-FEB-2001
LOCUS	AF318284 Drosophila melanogaster roughoid/rhomboid-3 mRNA, complete cds.
DEFINITION	AF318284
ACCESSION	AF318284
VERSION	AF318284.1 GI:12744729
KEYWORDS	
SOURCE	Drosophila melanogaster (fruit fly)
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridae; Drosophilidae; Drosophila. 1 (bases 1 to 1822) Wasserman,J.D., Urban,S. and Freeman,M. A family of rhomboid-like genes: Drosophila rhomboid-1 and roughoid/rhomboid-3 cooperate to activate EGF receptor signaling Genes Dev. 14 (11), 1651-1663 (2000) 20347042 JOURNAL MEDLINE PUBMED 10867159 REFERENCE 2 (bases 1 to 1822) AUTHORS Freeman,M., Urban,S. and Wasserman,J.D. TITLE Direct Submission JOURNAL Submitted (02-NOV-2000) Department of Cell Biology, MRC LMB, Hills Road, Cambridge CB2 2QH, UK Location/Qualifiers 1..1822 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /chromosome="3" /map="61F8" 326..1351 /note="contains 7 transmembrane domains" /codon_start=1 /product="roughoid/rhomboid-3" /protein_id="AAK06753.1" /db_xref="gi:12744730" /translation="MLLSGASQGRKRKNPLIGTVEEGMPSTVLQMPAPLSS SLTLLEVFILMWADPPEDSLLYVRPOQRQLMRFSLYALHLSWLHGYNVLTQL"
CDS	

BASE COUNT	449 a	482 c	456 g	435 t
ORIGIN				

	Query Match	Best Local Similarity	9.8%; Score 251.8; DB 3; Length 1822;
	Matches	413; Conservative	0; Mismatches 247; Indels 3; Gaps 1;
Qy	763	CCGTCGAGATTCGATGCTGATCGCATCGCCGACCGCGCTGTCGAGTGTGGCCCTCTTT	822
Db	683	CCAGAGGATTCCTCTTCGTGATATCGACCCGACGAGCGCTCGCAGCTCTGAGATTCCTT	742
Qy	823	AGCAACATGTTCCCTGCAACGCCAATCGATGATCACTCGGCGCTCAATATCGATCAGACTG	882
Db	743	AGCATAGCCCTGCTCCACGCCAGTTGAGCTCATCTGAGATTACANTGTGCTCACCTCAGCTC	802
Qy	883	TTTTTTGGCAATCCCTCGAGAGTGATGACACGGCACGGCCAGATCGGCGTGATCTACATG	942
Db	803	CTGTTCGGTGTGACCATTTGGAGCTGTACATGATTCGCTGAGAACGGGAGTCATTTATATG	862
Qy	943	GGGGGCGTTTTTGGCCGGAATCCCTGGGAGACCAAGTGTCTGCACTCGAGAGTCTTCCTGTG	1002
Db	863	GCCCGAATTTTGGCTGTGGCTCTTTGGGAACTTCGGTGTGGAATCGGAAATTTTCTGTG	922
Qy	1003	GGGCGACAGCGGTGCGCTATAGCCCTCTTGGCGGCATCTGGGCAACATTTACACTGAAC	1062
Db	923	GGAACTCTGTGAGAGATTTTACCTCTCTGTGGCCGCAACTGTGGAGATCTTTTGTCTAAC	982

Db	983	TTTGCTCAAAATGAGACACGCGGGTAAATTCAGCTATATGCTGATTTGTTGCTTTTCG	1042
Qy	1123	GATCTGGGCTATGCTCTCTACACCCAAATCTTGATGGAAGCGCTTCGCGAAGGTC	1182
Db	1043	GATTTGGGGATATGCGCTCTTATTCAGAGAACTGGCAATGCATCAGCTGCAGACAGACCC	1102
Qy	1183	CAGGTCGTGCATTTGCCACCTGACGGAGACCTGGCAGGACTCAACGATCGCTTTCTG	1242
Db	1103	TCTGTCTCTATATGCTCTACATATACCGGTGCTTTGGCCGGGAATTAGTGTGGTCTGCTG	1162
Qy	1243	GTGCTAAAGAACTTCGGTCAACCGAGATGAGAGCACTCATCTGTGGTCTAGCGTTGGGC	1302
Db	1163	CTTCTCCGGCAACTGGAATGTGTGGCTCGACACAGCCCACTCCGCTGATTGGCCCTGGGC	1222
Qy	1303	GTCCTACTGTGCTTCAACCGCTTTCGCCATCGTTTTCAACTGATCAACACGGTACCGCT	1362
Db	1223	GTATGTGTCATATTTACAGCGCTTTTGGATTCGCTTCAACTTGTGGAAATACGGTCAAGCT	1282
Qy	1363	CAGCTGATG---GAGAGACAGGGTGAAGATATTAACCAAGACTCGTTGACAGACTGGGA	1419
Db	1283	CAACTTTCGGCCGAGAGAGAGGGCCAGGTATCATCGACAGCACTGTGATGAATGACCTGGGA	1342
Qy	1420	GTG 1422	
Db	1343	ATG 1345	

RESULT 10	
AF318283	
LOCUS	2442 bp
DEFINITION	Drosophila melanogaster rhomboid-2 mRNA, complete cds.
ACCESSION	AF318283
VERSION	AF318283.1
KEYWORDS	GI:12744727
SOURCE	
ORGANISM	Drosophila melanogaster (fruit fly)
	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
	Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 2442)

AUTHORS Wasserman, J.D., Urban, S. and Freeman, M.
 TITLE A family of rhomboid-like genes: Drosophila rhomboid-1 and
 rhomboid-3 cooperate to activate EGF receptor signaling
 JOURNAL Genes Dev. 14 (13), 1651-1663 (2000)
 MEDLINE 20347042
 PUBMED 10867155
 REFERENCE 2 (bases 1 to 2442)
 AUTHORS Freeman, M., Urban, S. and Wasserman, J.D.
 TITLE Direct Submission
 JOURNAL Submitted (02-NOV-2000) Department of Cell Biology, MRC LMB, Hills
 Road, Cambridge CB2 2QH, UK
 FEATURES
 source
 1. 2442
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /chromosome="3"
 /map="62A1"
 388..1683
 /note="contains 7 transmembrane domains"
 /product="rhomboid-2"
 /codon_start=1
 /protein_id="AAK06752.1"
 /db_xref="GI:1274728"
 /translation="MRVQCTNDAGYDTEHTSINSDEAELELRRELLRDKMLFDMFD
 PEGRFISVEEFLVALKSEPEFLSQVPMNRKELLEPAKAKLPAAGGYTFODPVVM
 SGKRTSPKCAVHNRDREYSENPOLVPLPRKHAVALMELPERDKXYAD
 RYTCPPRPFLITLVGLFPRVHVSVMGEAPRPIISDSMFIRPKRHEIMFL
 FTMVLAHGLHGFNVAVOLVGLPLDMVHGSTRICITSGVLSLSTISFDPDV
 LVGASGVYALAAHLANVLNVLHQRVAVIKLHLVPSDFPGAVIARVAGDGLQ
 LGSSEBPFIADLETGAVSVYAHLAGALVGLTIGLVKSPFQRLHQLDLMWIALGT
 YVALVFAIFNINMGPMFNIRVEKIRVETIFNDPQV"
 BASE COUNT 626 a 568 c 593 g 655 t
 ORIGIN
 Query Match 7.7%; Score 198.2; DB 3; Length 2442;
 Best Local Similarity 56.0%; Pred. No. 1.5e-29;
 Matches 442; Conservative 0; Mismatches 308; Indels 39; Gaps 2;
 QY 600 GGCACGACCATCTCCGAGTCCGAGATATCGCCCTGAGTACGTCACCGCCAGCA 659
 DB 819 GGCACGAGATCTCTCCGAGAGGGGAGCCGCAAGTACTACGCGATCGCTACACCTG 878
 QY 660 CTGGCCCTGTCATCTGATCTCCATCATCTGATGATGACATCTTGCGCTAGACCG 719
 DB 879 CTGTCCGCGCCCTTTTTCATCATCTGCTGCTGAGAGCTGGGTTTCTTTGTTTA 938
 QY 720 CTACACATGCGCCGACGAGATTTCCGAGTACCCGTTCCGATTCGTCGATTCGGTCT 779
 DB 939 CCACTCAGTGTGACGAGGTGAGCGCGGCGCAAGGGAGCCCATACCTCGGACTCCATGTT 998
 QY 780 GGTCTATCGCGGAGACCGCGCTGTGAGGTGCGCTCTTTAGTACATGTTCTCTGCA 839
 DB 999 CATCTACCTCCCGATAGGCTGACAGATCTGCGCTTCTCTTACATGCTGCTCCA 1058
 QY 840 CGGCACTGGTTCACCTGGGCTTCAATATTCGATCCAGCTGTTTGGATTCCTCCCT 899
 DB 1059 TGCCTGGGTGCTGATCTGGTTTAACTGGGCACTTCACTGCTTTTGGCTGCACT 1118
 QY 900 GGAAGTATGACGAGCAGCGCCAGATCGGCTGATCTACATGCGCGGCGTTTTCGCG 959
 DB 1119 GGAAGATGTTCAATGTTTCAAGAGATAGCTGATCTACTTCTCCGGGGTGTGCGGG 1178
 QY 960 ATTCCTGGGACCACTGTGTGTGACTCGAGAGTCTTCTGCTGGGCGCCAGCGGTGCGT 1019
 DB 1179 AAGCCCTGGGAACATGATCTTTTGAACCCGAGTGTCTTGTGTGGGCGCCAGTGTGAGGT 1238
 QY 1020 CTATGCGCTGTTGGCGGACATCTGGCAACATTTACATGAACTATAGCGGACATGAAG 1079
 DB 1239 ATTAGCTTTGCTAGCGCCCACTGGCAATGCTCTTCACTATACCAAGATGGGTTA 1298
 QY 1080 CGCATCCACGCACTCGGATCCGTTGTCACTTTGTCTCTCGATCTGGGCTATGCTCT 1139

DB 1299 TGGAGTCAATCAAGCTGCTTCAATCTGTGCTTTCCTTCCTGCACTTTGGCTTGTAT 1358
 QY 1140 CTACACCCATA-----CTTCGATGAAGCGCTTC--- 1170
 DB 1359 CTATGCTCCCTATGCTGGGAGCAACTGACAGCTGGATCTCTGAGTGAATTTCCGGCCAT 1418
 QY 1171 -----GCCAAGGTTCCCAAGTGTGTATCATTTGCCCATGACGGAGCCCTGCG 1220
 DB 1419 AGATCAGCGGAAACGGCGGGCGCTGTTGATGTGGCCATTTGGCTGTGCCATTGCG 1478
 QY 1221 AGACTTAAGATTCGCTTTCTGTGTAAAGACTTCGTCACCGAGATGACAGAGCT 1280
 DB 1479 TGGCTGATATCGGGCTCTGTGCTCAAGCTTTGACAGAGGCTCCACAGACGCT 1538
 QY 1281 CATCTGTGGCTAGCGGTGGAGGCTACTGTGCTTCACTGCTTCCAGCTTCCTGCTTTTCA 1340
 DB 1539 GCTTGTGTGATCGCTTGGGAACGTACTGGCGCTGTGTGCTTCCCATCGCTTCA 1598
 QY 1341 CCTGATCAA 1349
 DB 1599 CATCATGAA 1607
 RESULT 11
 AY119150
 LOCUS AY119150 1900 bp mRNA linear INV 16-JUN-2002
 DEFINITION Drosophila melanogaster RB60377 full insert cDNA.
 ACCESSION AY119150
 VERSION AY119150.1 GI:21430663
 KEYWORDS FLI, CDNA.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1900)
 Stapleton, M., Broksstein, P., Hong, L., Agbayani, A., Carlson, J.,
 Champe, M., Chavez, C., Dorsett, V., Dreene, D., Farfan, D., Fries, B.,
 George, R., Gonzalez, M., Guarin, H., Kronmiller, B., Li, P., Liao, G.,
 Miranda, A., Mungall, C.J., Nunoo, J., Pacled, J., Paragas, V., Park, S.,
 Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Rubin, G.M.
 and Celniker, S.
 Direct Submission
 Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,
 Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, USA
 COMMENT
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This clone was sequenced as part of a high-throughput process to
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
 Science 2000). The sequence has been subjected to integrity checks
 for sequence accuracy. Presence of a polyA tail and contiguity
 within 100 kb in the genome. Thus we believe the sequence to
 reflect accurately this particular cDNA clone. However, there are
 artifacts associated with the generation of cDNA clones that may
 have not been detected in our initial analyses such as internal
 priming, priming from contaminating genomic DNA, retained introns
 due to reverse transcription of unsplined precursor RNAs, and
 reverse transcriptase errors that result in single base changes.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our web site
 (http://fruitfly.berkeley.edu) or send email to
 cdna@fruitfly.berkeley.edu.
 FEATURES
 source
 1. 1900
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /strain="y; cn bw bp"
 /db_xref="taxon:7227"
 1..1900
 /gene="rho-2"

CDS

/db_xref="FLYBASE:FBgn0041725"
 158..1150
 /gene="rho-2"
 /note="Longest ORF"
 /codon_start=1
 /product="Rb60377p"
 /protein_id="AA051010.1"
 /db_xref="GI:21430664"
 /translation="MSGKTRSPKCAVHHDRVCSNDPOLVINEPPLFRKMHVA
 MEILPBRDKVYADRYTCPPPEFIIIVLVGPNVSVVTVGAAPRGPISDSM
 FLYRPRKREHTRFLPYVTHAGMHLGKRVANQVLFGLPLMWHSTRACIYFSGV
 LAGSLGTSIFDPVFLVAGSGVYALIAHLANVLNHYQMRGVYIKLHLIVFSD
 FGPAIYARVAGDELQUGSSSEFLAIDQAEAGAVSVYVHLAGIAGLTGLVKSSE
 OKHLEOLMIALGTLYALVPAIFAFINMGFAMFNIRVEKIRVETETIFNDPQV"

BASE COUNT 500 a 391 c 433 g 576 t

ORIGIN

Query Match 7.7%; Score 197; DB 3; Length 1900;
 Best Local Similarity 58.3%; Pred. No. 2.5e-29;
 Matches 390; Conservative 0; Mismatches 240; Indels 39; Gaps 1;

720 CTACACAAATGCCCGCCAGAAATTTGGGGCTAACCCGTTCCGATTCGTCGATTCGGTCT 779
 406 CCACTCAGTGTGTACGGGTAGAGCGCGCCAGGAGCCATACCTCGACATTCATGTT 465
 780 GGTCTATCGCGGAGACCGGCGTCTGACAGTGTGGCGCTTCTTATGATCAATGTTCTGCA 839
 466 CATCTACCTCCGATTAAGCTCAGAGATCTGGCGCTTCTCTTCAATGATGATCCCA 525
 840 CGCCAACTGGTTCACCTGGGCTTCAATATGTCATCACTGTTTGGCATTTCCCT 899
 526 TGCAGGATGCTCATCTCGGTTTAACTGGGCGATTCACTGGTTTGGGCTGCGACT 585
 900 GGAAGTATGACAGGAGCGGCGGATCGGCGATGATCAATGACGGCGCTTTTGGCGG 959
 586 GGAAGTATGATCAATGTTTCAAGAGATAGCTGATCTTCTCCGAGGATGCTGGCGG 645
 960 ATCCCTGGGACCAAGTGTGTCAGTGTGAGAGTCTTCTGATGGGCGCCAGCGGTGCGT 1019
 646 AAGCTTGGGACATGATCTTTGACCCGATGTGTTCTTGTGGGCGCCAGTGTGAGT 705
 1020 CTATGCCCTTGGCGCGCACATCTGGCCAACTTAACTGAATGCGGACATGAGAG 1079
 706 ATAGCTTTGCTAGCGCGCCACCGGCAATGCTTCTCACTATCACCAGATGCGTTA 765
 1080 CGCATCCAGCAACTCGGATCCGTTGATCTTGTCTCTCGGATCTGGGCTATGCTCT 1139
 766 TGGAGTCATCAAGCTGCTTCAATCTGATCTTCTGATCTTCTGCACTTGGCTTGGCAT 825
 1140 CTACACCC-----ATACTTCGATGG 1160
 826 CTATGCTGCTATGCTGGCGACGATGCGAGTGGAGTATCAAGTGAATTTCTTGCGCAT 885
 1161 AAGCGCTTTGCCAAGGAGTCCCAAGTGTGTCATTTGCCACCTGACGGAGCGCTGCG 1220
 886 AGATCAGCGGAGAAACGGCGGGTGTCTTTGTATGTCGCCCATTTGGCTGGTCATTCG 945
 1221 AGACATACGATGGGCTTTCTGTGCTTAAAGAACTTGGTCAACGAGATACAGAGCGCT 1280
 946 TGGCTTAACATATCGGGCTGTGCTCAAGAGCTTGTGAACAGAGCTTCAACGAGCGCT 1005
 1281 CATCTGGTGGCTAGCGTGGGCGCTCTACTGTCCTTCAACGCTTCTTGGCAATCGTTTCAA 1340
 1006 GCTCTGGTGGATGCTTGGGAGACGTACTTGGCGCTGGTGTCTTCCCATTCGCTTCAA 1065
 1341 CCTGATCAA 1349
 1066 CATCATGAA 1074

LOCUS BT003499 2107 bp mRNA linear INV 12-FEB-2003
 DEFINITION Drosophila melanogaster RE43527 full insert cDNA.
 ACCESSION BT003499
 VERSION BT003499.1 GI:28316963
 KEYWORDS
 SOURCE
 ORGANISM
 Drosophila melanogaster (fruit fly)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephyroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 2107)
 Stapleton,M., Broksstein,P., Hong,L., Agbayani,A., Carlson,J.,
 Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frieze,E.,
 George,R., Gonzalez,M., Guarin,H., Kronmiller,B., Li,P., Liao,G.,
 Miranda,A., Mungall,C.J., Nuno,J., Paclob,J., Paragas,V., Park,S.,
 Patel,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
 and Celniker,S.
 Direct Submission
 Submitted (12-FEB-2003) Berkeley Drosophila Genome Project,
 Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, USA
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This clone was sequenced as part of a high-throughput process to
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
 Science 2000). The sequence has been subjected to integrity checks
 for sequence accuracy, presence of a polyA tail and contiguity
 within 100 kb in the genome. Thus we believe the sequence to
 reflect accurately this particular cDNA clone. However, there are
 artifacts associated with the generation of cDNA clones that may
 have not been detected in our initial analyses such as internal
 priming, priming from contaminating genomic DNA, retained introns
 due to reverse transcription of unspliced precursor RNAs, and
 reverse transcriptase errors that result in single base changes.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our web site
 (<http://fruitfly.berkeley.edu>) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES
 source
 1..2107
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /strain="Y; con bw sp"
 /db_xref="taxon:7227"
 1..2107
 /gene="rho-2"
 /db_xref="FLYBASE:FBgn0041725"
 62..1357
 /gene="rho-2"
 /note="Longest ORF"
 /codon_start=1
 /product="Rb3527p"
 /protein_id="AA039503.1"
 /db_xref="GI:28316964"
 /translation="MRVCTNDAGVDTHTSTNSDPEALRELLRDKMLLFDMFD
 PEGRGISVEEPELALKSREPLISQVPMNRRELLRBAKAKLPTGGVYTFDDFVVM
 SGRKTRSFCAVHHDRVCSNDPOLVINEPPLFRKMHVAMEILPBRDKRYAD
 RYTCPPPEFIIIVLVGPFVYHVSIVGEAARPIPSDSKFIYRDKRHIMFL
 FVVLHAGMHLGKRVANQVLFGLPLMWHSTRACIYFSGVLASLGTSIFDPVFL
 LVAGSGVYALIAHLANVLNHYQMRGVYIKLHLIVFSDFGPAIYARVAGDELQ
 LGSSEFLAIDQAEAGAVSVYVHLAGIAGLTGLVKSSEFGKLEOLMIALGT
 YIALVPAIFAFINMGFAMFNIRVEKIRVETETIFNDPQV"

BASE COUNT 539 a 474 c 496 g 598 t

ORIGIN

Query Match 7.7%; Score 197; DB 3; Length 2107;
 Best Local Similarity 58.3%; Pred. No. 2.6e-29;
 Matches 390; Conservative 0; Mismatches 240; Indels 39; Gaps 1;

720 CTACACAAATGCCCGCCAGAAATTTGGGGCTAACCCGTTCCGATTCGTCGATTCGGTCT 779

```

Db      613  CCACCTCAGAGTGCACCGGTGAGCGCGCCGCAAGGGGAGCCATACCTCCGACTCCATGTT 672
Qy      780  GGTATATCGGCGGAGCCGCGCTCTGAGGAGTGGCGCTCTTAGTACATGTTCCCTGA 839
Db      673  CATCTACCTCCCGATAGCGATCGAGATCTGGCCCTTCTCTTACATGATGACTCCA 732
Qy      840  CGCCAACTGGTTCACCTGGGCTTCAATATCGTATCATCGACTGTTTGGCATTTCCCT 899
Db      733  TCCCGGGTGGCTGACCTCGGTTTAACTGGAGTTGACGCTGTTTGGGCTGCACT 792
Qy      900  GGAAGTATGACAGCGACGCGAGATCGGCGTATCTACATGCGCGGCGTTTCCCG 959
Db      793  GGAAGTATGACAGCGAGATCGGCGTATCTCTCCGCGGCTGTCGCGG 852
Qy      960  ATCCCTGGGACACAGTGTGAGCTCGGAGGCTCTCGTGGGCGGACGCGGTGCGCT 1019
Db      853  AAGCTGGGACATGATCTTGAAGCCGAGATGTCTGTGGGCGGACGAGTGGAGT 912
Qy      1020  CTATGCTCTGTTGGCGGACATCTGGCAATTAACATGAATGACGACATGAAGAG 1079
Db      913  ATACGCTTGTAGCGGCCCACTGGCAATGTCTTCACTACATACACAGATGGCTTA 972
Qy      1080  CGCATTCACGCAACTCGATCCGTTGTATCTTGTCTCGGATCTGGGCTATGCTCT 1139
Db      973  TGGAGTATCAAGCTGCTTCAATCTGGCTTCTGTTCTTCACTTGGCTTTCGAT 1032
Qy      1140  CTACACCC-----ATACTTGATAG 1160
Db      1033  CTATGCTCTGATATGCTGGGAGAGATCGAGCTGGATATCGAGTAAATTCCTGGCAT 1092
Qy      1161  AAGCGCTTGGCGCAAGGATCCCGAGGTGCTGATATTTGCCCACTGACGCGAGCCCTG 1220
Db      1093  AGATACGAGGGAACGCGCGGCTGTCTGTATGTGGCCCATTTGGCTGGTGCATTC 1152
Qy      1221  AGGACTAACGATCGGCTTCTGTGCTAAAGAACTTGGCTACCGAGATGACGAGCT 1280
Db      1153  TGGCTTAATCTATCGGCTGCTGTGTCTCAAGAGCTTGAAGAGACTCCAGAGAGCT 1212
Qy      1281  CATCTGTGGCTAGCTTGGGCGCTCTACTGTGCTTACCGCTTTCGCAATGTTTCA 1340
Db      1213  GCTTGTGTGATCGCTTGGGAACGTATCTGGCTGTGTCTTGCATGCGCTTCA 1272
Qy      1341  CCTGATCA 1349
Db      1273  CATCATGAA 1281

RESULT 13
AC017712      35808 bp      DNA      linear      HTG 10-DEC-1999
LOCUS      AC017712
DEFINITION      Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION      AC017712.1 GI:6554285
VERSION      HTG: HTGS PHASE2.
KEYWORDS      Drosophila melanogaster (fruit fly)
SOURCE      Drosophila melanogaster
ORGANISM      Drosophila melanogaster
REFERENCE      Adams M. and Venter, J.C.
AUTHORS      Direct Submission
TITLE      Submitted (10-DEC-1999) Celera Genomics, 45 West Gude Drive,
JOURNAL      Rockville, MD, USA
COMMENT      This sequence was identified as CDM:10211825 by the submitter.
      For more information on this record e-mail to fly@celera.com.
      * NOTE: This is a 'working draft' sequence.
      * This sequence will be replaced
      * by the finished sequence as soon as it is available and
      * the accession number will be preserved.
FEATURES
      source
      1..35808

```

```

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
BASE COUNT      10107 a 7885 c 7738 g 10078 t
ORIGIN
Query Match      6.7%; Score 170.8; DB 2; Length 35808;
Best Local Similarity 56.3%; Pred. No. 9.8e-24;
Matches 413; Conservative 0; Mismatches 247; Indels 74; Gaps 2;

Qy      763  CCGTGGATTCGGTCTGTATGAGCGGACCGGCTCTGAGTGTGGGCTTCTT 822
Db      23650  CCAGAGATTCCTTCTGTATATGACCCGACAGCGCTTACGAGATTCCTT 23709
Qy      823  AGCTACATGTTCTGACAGCCAACTGTTTCCACTGGGCTTCAATATGCTATCCAGCTG 882
Db      23710  AGCTATGCCCTCTCAAGCCAGTTGGCTTCACTTGGGTTAATATGCTCACTAGCTC 23769
Qy      883  TTTTGTGGCATTCCTGTGGAGTGAATGACAGGACGCGGAGATCGGCTGATCTACATG 942
Db      23770  CTGTTGGTGTGCTATGAGCTGGTGAATGATGCTGATGAGAGAGGATCATTTATATG 23829
Qy      943  GCGGCGTTTGTGCGGATCCTTGGGACAGTGTGCTGACTCGAGGTTCTTCTGGTG 1002
Db      23830  GCGGAGTTTGTGCTGCTTGTGGGAACTTGGTGTGAGTCTCGAAGTTTCTGCTG 23889
Qy      1003  GGGGCGAGGAGGAGGCTATGAGCTTGGGCGGACATGAGGCAATTAACATGAGAC 1062
Db      23890  GAGGCGAGTGAAGAGTTTACCTTGTGGCGGCGGCACTGGGCTTGTGCTCAAC 23949
Qy      1063  TATGCGCATGAAAGGCGATCCAGCACTCGGATCCGTTGTATCTT----- 1112
Db      23950  TTTGTCAATATGAGACAGGAGTATTAACATATGAGCTGATTTTGTGGTAAATGA 24009
Qy      1113  ----- 1112
Db      24010  ACTGTATATATCTATATATGAAATTTGCTGTAGAAATCAATTAATTTCTATCTTAA 24069
Qy      1113  -TGCTCCGAGATCGGGCTATAGCTCTCAACCAATCTTGAATGAGAGCGCTTG 1171
Db      24070  CAGTTTTCGATTTGGGATAGGCTCTTATTAAGAACTGGCAATGATGATGAGCTGC 24129
Qy      1172  CCAGGGTCCCGAGGTGCTGATATTTGCTTGGCCACTGAGGAGCCCTGGCAGACTTAACA 1231
Db      24130  AAGACAGACCTCTGTCTCTATATGCTGATACATACCGGTCTCTGGCGGAGTTAAGT 24189
Qy      1232  TCGGCTTCTGTGCTTAAAGAACTTGGCTACCGAGATTAAGAGACTCATCTGGTGC 1291
Db      24190  TGGTCTGTGCTTCTCGGCAACTGATGATGGGCTGGCAACGCGCACTCCGCTGGT 24249
Qy      1292  TAGCGTTGGGCTCTACTGTGCTTCAACCGTCTTGGCATCGTTTCAACCTGATCAACA 1351
Db      24250  TGGCCCTGGGCTATGATGATATTCAGCGCTTGGATTCGCTTCAACTTGGGAATA 24309
Qy      1352  CGGTACCGCTCAGCTGATG---GAGGACAGAGGTGAGTGAATTAACCGCATCTGTTGC 1408
Db      24310  CGGTACAGCTCAATCTTGTGGCGGAGAGAGAGGCGCAGGTATCAAGGACATCTGATTA 24369
Qy      1409  AGCACTGGGAGTG 1422
Db      24370  ATGACCTCGGAATG 24383

RESULT 14
AC010007      175481 bp      DNA      linear      INV 01-JUN-2002
LOCUS      AC010007
DEFINITION      Drosophila melanogaster 3L BAC RP98-17K17 (Ruswell Park Cancer
ACCESSION      AC010007 GI:21306589
VERSION      HTG.
KEYWORDS      Drosophila melanogaster (fruit fly)
SOURCE      Drosophila melanogaster
ORGANISM      Drosophila melanogaster

```

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE AUTHORS

1 (bases 1 to 175481)
Muzny, D., Scherer, S., Adams, M.D., Holt, R.A., Evans, C.A., Gocayne, J.D., Tabor, P., Williamson, A., Hong, F.H., Dugan-Rocha, S.D., Sodergren, E.S., Hodgson, A.R., Chen, R.C., Ayele, M., Scott, G.S., Worley, K.W., Amaratunga, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Beeson, K.Y., Brown, M., Buhay, C., Busam, D.A., Center, A., Chen, G., Chen, Z., Clerc-Blankenburg, K., Davenport, L.B., Dietz, S.M., Ding, Y., Dodson, K., Doup, L.E., Drepper, H., Emery-Cohen, A., Fertire, S., Garg, N.D.S., Houck, J.J., Hostin, D., Howland, T.J., Hume, J., Ibegwam, C., Jalali, M., Kovac, C., Liu, M., Mattei, B., McIntosh, T.C., Morgan, M., Moy, M., Murphy, B., Nelson, K.A., Ndassa, Y., Nguyen, N., Perez, L., Pittman, G.S., Pui, Y., Scheeler, F., Shen, H., Strong, R., Tector, C., Wang, Q., Williams, S.M., Xiang, J., Zaveri, U.S., Zhou, J., Zorrilla, S., Smith, H.O., Wheeler, D., Weinstock, G., Gibbs, R. and Venter, J.C.

TITLE

Direct Submission

JOURNAL REFERENCE AUTHORS

2 (bases 1 to 175481)
Worley, K.C., Adams, C., Adio-Oduola, B., Ali-oman, F.R., Allen, C., Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Banks, T., Barbata, J., Benton, J., Bimaye, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowles, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.J., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dachtore, S.R., David, R., Davila, M.L., Davis, C., Denny, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earmhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Guarnatone, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homel, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudh, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovac, C., Kratoch, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, Z., Li, Z., Lichtrage, O., Lien, C., Liu, J., Liu, W., Lounsbeg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkwo, S., Ogih, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Stinson, I., Sodergren, E., Sonalike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, K., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wiczek, R., Woodson, S., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.

TITLE

Direct Submission

JOURNAL

Submitted (29-DEC-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

BCM-HGSC.

AUTHORS

Submitted (01-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

TITLE

On Jun 1, 2002 this sequence version replaced g1:1798562.

JOURNAL

Submitted (01-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jun 1, 2002 this sequence version replaced g1:1798562.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for

a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

```
source
  1.175481
  /organism="Drosophila melanogaster"
  /mol_type="genomic DNA"
  /db_xref="taxon:7227"
  /chromosome="3L"
  /clone="RP98-17K17"

misc_feature
  22738
  /note="Calling a C and should be T"
  /function="polymorphic site"

misc_feature
  23469
  /note="Calling a C and should be T"
  /function="polymorphic site"

misc_feature
  23528
  /note="Calling a C and should be T"
  /function="polymorphic site"

misc_feature
  23693
  /note="Calling a C and should be T"
  /function="polymorphic site"

misc_feature
  23870
  /note="Calling an A and should be G"
  /function="polymorphic site"

misc_feature
  23932
  /note="Calling a G and should be A"
  /function="polymorphic site"

misc_feature
  24185
  /note="Calling a C and should be T"
  /function="polymorphic site"

misc_feature
  24507
  /note="Calling a C and should be T"
  /function="polymorphic site"

misc_feature
  24576
  /note="Calling a C and should be T"
  /function="polymorphic site"

misc_feature
  24936
  /note="Calling a C and should be T"
  /function="polymorphic site"

Query Match
  6.6%; Score 169.8; DB 3; Length 175481;
  Best Local Similarity 56.2%; Pred. No. 2.2e-23;
  Matches 413; Conservative 0; Mismatches 247; Indels 75; Gaps 2;

763 CCGTGGATTCGGTCTGCTGATCGCGCGAGCGGCTGCGAGTGGCGCTTTCTTT 822
  |||
166876 CCAGAGGATTCCTCTCTGATATACACCCGACACCGCGCTGAGATTTCTTT 166935
  |||

823 AACTCATTTCTCTGAGCGCAACTGCTTCCACCTGGGCTTCAATTCGTCATCCAGCTG 882
  |||
166936 AACTATGCCCTGCTCAACCGCATGTCGTCATCTGGGTTACATGCTCACTCAGCTC 166995
  |||

883 TTTTGGCATTCCTCTGAGGATGATGACGCGACGCGAGATCGGCTGATCTTACATG 942
  |||
166996 CTGTTGCGTGTGCGCATGAGAGCTGATCATGATCGCTGAGAAAGGAGATCTTATATG 167055
  |||

943 GCGGCGTTTTCGCGGATTCCTGGGACCAAGTGTGTCGACTCGGAGGCTTCTCTGTTG 1002
  |||
167056 GCGGAGTTTTCGCTGCTCTTTGGGAACTTCGCTGCTGAGACTCGGAACTTTTCTGTTG 167115
  |||

1003 GCGGCGAGGCTGAGGCTATGCGCTTATGCGCGACATCTGGCGCAACTTACATGAGAC 1062
  |||
167116 GAGGCGAGTGAAGAGATTTACGCTTGTGCGCGCCCAACTGCGAGTCTTTTGTCTCAAC 167175
  |||

1063 TATGGGCATGAAAGCGCATTCACGCACTCGATTCGCTTGTCTATCTT----- 1112
  |||
```

```
Db 167116 TTTGGTCAATGAGACACGGGGTAAATAGCTAAATGCTGTGATTTTGTGGTAAATGA 167235
Qy 1113 ----- 1112

Db 167236 ACTGGTAAATTTCTAATATGAAATTTGCTGTAGAAATCATAAACATTTTCTATCTTT 167285
Qy 1113 --TGTCTCTGCGATCTGGGCTATGCTCTCTTACACCATATCTGATGAAAGCGCTTC 1170
Db 167296 ACAGTTTCTGCGATTTGGGGATGATGCCCTCTATTCAAGAAATCGGCATGATCAGCTG 167355
Qy 1171 GCCAAGGTTCCCAAGTGTGTCATTTGCCACCTGACGGGAGCCCTGGCAGACTAACG 1230
Db 167356 CAGACACGACCTCTGTCTCTCTATATGCTCAACATGACCGGTGCTTGGCCGGATTACT 167415
Qy 1231 ATCGGCTTTCTGATGCTAAAGAACTTCGGTCAACCGAGATAGAGAGCTCATCTGTTGG 1290
Db 167416 GTGGTCTGCTCTTCTTCCGGCAACTGATGATGGGCTGCGACACGCCACTCGCTGG 167475
Qy 1291 CTAGGTTGGGCGTCTACTGTGCTTCAACCGCTTTCGCCATGCTTTTCAACCTGATCAAC 1350
Db 167476 TTGGCCCTGGGGGTATGATGATTCATTCAGCGCTTTTGAATGCGCTTCAACTTGGTGAAT 167535
Qy 1351 ACGTGACCGCTCACTGATG--GAGAGACAGGCTGATGATTAACCATCTGTTG 1407
Db 167536 ACGGTACAGCTCAACTTCTGGCCGAGAGAGGCGCAGGTGATGACGACATCTGATG 167595
Qy 1408 CACGACCTGGGAGTG 1422
Db 167596 AATGACCTCGGAATG 167610

RESULT 15
AC005847/c 268369 bp DNA linear INV 22-OCT-1998
LOCUS
DEFINITION
  AC005847 Drosophila melanogaster DNA sequence (Plus DS03179 (D226), DS06357
  (D230), DS03404 (D231), DS06962 (D232), and DS07291 (D240)),
  complete sequence.
ACCESSION
  AC005847 AC004331 AC004436 AC004434 AC004488 AC004489
VERSION
  AC005847.1 GI:3779012
KEYWORDS
  HTG.
SOURCE
  Drosophila melanogaster (fruit fly)
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 268369)
  Celniker,S.E., George,R.A., Galle,R.F., Hoskins,R.A.,
  Svitzkas,R.R., Harris,N.L., Agbayani,A., Arcaina,T.T., Baxter,E.,
  Blazej,R.G., Chavez,C., Chew,M., Doyle,C.M., Farfan,D.E.,
  Flanagan,J., Houston,K.A., Hummasti,S.R., Katta,K., Kearney,L.,
  Kim,S.H., Lee,B., Lomocan,M.A., Mak,J., Mazda,P., Moshrefi,A.R.,
  Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Punch,E.,
  Smit,E., Twomey,B., Wan,K.H., Zhang,R., Zieran,L.L. and Rubin,G.M.
  Sequencing of Drosophila chromosome 3L, region 61F3-62A2
  Unpublished (1997)
  2 (bases 1 to 268369)
  Celniker,S.E., George,R.A., Galle,R.F., Hoskins,R.A.,
  Svitzkas,R.R., Harris,N.L., Agbayani,A., Arcaina,T.T., Baxter,E.,
  Blazej,R.G., Chavez,C., Chew,M., Doyle,C.M., Farfan,D.E.,
  Flanagan,J., Houston,K.A., Hummasti,S.R., Katta,K., Kearney,L.,
  Kim,S.H., Lee,B., Lomocan,M.A., Mak,J., Mazda,P., Moshrefi,A.R.,
  Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Punch,E.,
  Smit,E., Twomey,B., Wan,K.H., Zhang,R., Zieran,L.L. and Rubin,G.M.
  Direct submission
  Submitted (22-OCT-1998) Berkeley Drosophila Genome Project, MS
  64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
  Berkeley, CA 94720, US
  Sequence submitted by:
  Berkeley Drosophila Genome Project
  Lawrence Berkeley National Laboratory, MS 64-121
  Berkeley, CA 94720
  For further information about this sequence, including its location
  and relationship to other sequences, please visit our sequence
```

archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.
Library locations: 11-34, 21-67, 44-36, 73-50, 92-76.
Location/Qualifiers

FEATURES

Source

```
1. 268369
/organism="Drosophila melanogaster"
/mol_type="Genomic DNA"
/db_xref="taxon:7227"
/chromosome="3L"
/map="61f3-62A2"
/clone="P18 DS03179 (D226), DS06357 (D230), DS03404
(D231), DS06962 (D232), and DS07291 (D240)"
/note="These five P1s were completed as a project. The
coordinates are the P1 ends for each P1. DS03404
(D231) extends from bp 1 to bp 80,732. DS03179 (D226)
extends from bp 32,734 to bp 117,954. DS06357 (D230)
extends from bp 88,745 to bp 169,482. DS06962 (D232)
extends from bp 127,117 to bp 209,427. DS07291 (D240)
extends from bp 194,759 to bp 268,369."
```

BASE COUNT 71701 a 61029 c 61736 g 73903 t
ORIGIN

Query Match 6.6%; Score 169.8; DB 3; Length 268369;

Best Local Similarity 56.2%; Pred. No. 2.4e-23;

Matches 413; Conservative 0; Mismatches 247; Indels 75; Gaps 2;

```
QY 763 CCGTCGGATTGGGTCGTCTATCGGCGGACCGGCGTGCAGAGTGGCGCTTCTT 822
DB 58790 CCAGAGGATCTCTTCGTATATCGACCGACGCGCCTCGAGCTCTGAGATTCCT 58731
QY 823 AGCTACATGTTCTGCAAGCCAACTGGTTCACCTGGGCTTCAATATGTCATCCAGCTG 882
DB 58730 AGCTATGCCCTGCTCCACGCGCATGTGGCTCATCTGGGTTACAATGCTCAGCTCAGCTC 58671
QY 883 TTTTTCGATTCCTCCCTGAGGTGATGACGCGCACGCGCATCGGCGGTGATCTACATG 942
DB 58670 CTGTCGGTGTGCATTTGAGCTGTGATCATGATTCGCTGAGAACGCGAGTCATTTATATG 58611
QY 943 GCGGGGGCTTTTCCCGGATCCCTGGGGACACAGTGTGTGACATCGAGAGTCTTCCTGGTG 1002
DB 58610 GCGGAGTTTGTGCTCTTTGGGAACTTCGGTGTGAGCTCGGAACTTTTCTGTGTG 58551
QY 1003 GCGCGCAGCGGTGGCTCTATGCGCTTGTGGCGCACATCTGGCCAACTTATACATGAAAC 1062
DB 58550 GGAGCCAGTGGAGAGTTTACGCTTGTGCGCGCCCACTGGGAGTCTTTGTCTCAAC 58491
QY 1063 TATGCGCACATGAAGCGCATCCACGAACTGGATCGTTGTCTT----- 1112
DB 58490 TTGTGTCAAATGAGACACGGGGTAAATCAGCTAATGCTGTGATTTGTTCGGTAACTGA 58431
QY 1113 ----- 1112
DB 58430 ACTGATTAATTTCTAATATGAATTTGCTTAGGAATCATMAAAATTTTCTATCTCTT 58371
QY 1113 --TGTCCTGCGATCTGGGCTATGCTCTACACCCAAATCTTGATGGAAGCGCCTTC 1170
DB 58370 ACAGTTTCTGCGATTTGGGGTATGCCCTCTATTCAGAGAACTGGCAATGCAATCAGCTG 58311
QY 1171 GCCAAGGTTCCCAAGTGTGTACATTTGCCACTGACGGAGCTCTGGCAGAGCTAAGC 1230
DB 58310 CAGACACGACCTGTCTCTAATATGCTCACATGACCGGTGCTTGGCCGGGATTAAGT 58251
QY 1231 ATGGGCTTTTGTGCTTAAAGAACTTGGTCAACGAAAGTACAGAGCTCATCTGTGG 1290
DB 58250 GTGGGTCTGTCTCTCCGCAACTGAGTGTGGGCTGCGACACGCCCACTCCGCTGG 58191
QY 1291 CTAGCGTTGGGCGCTCTACTGCTTACCGCTTTCGCCATCTGTTTCAACTGATCAAC 1350
DB 58190 TTGGCCCTTGGGCGGATGTGATTTCAAGCCCTTTGGAATCGCCTTCAACTTGTGTAAT 58131
QY 1351 ACGGTGACCGCTCAGCTGATG--GAGAGACAGGAGTGAATTAACCAAGCATCTGTTG 1407
DB 58130 ACGGTCAACGATCTTGTGGCCGAGAGAGGAGGCGAGGTGATCAGGACATCTGATG 58071
```

QY 1408 CACGACCTGGAGATG 1422
DB 58070 AATGACCTGGAGATG 58056

Search completed: February 8, 2004, 04:40:50
Job time: 9761.18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:33:25 ; Search time 728.807 Seconds
(without alignments)
9474.605 Million cell updates/sec

Title: US-09-614-150A-8

Perfect score: 2558

Sequence: 1 ggcgcgcgccttccttc.....taatacaaatcaagca 2558

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneq/geneq-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneq/geneq-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneq/geneq-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneq/geneq-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneq/geneq-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneq/geneq-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneq/geneq-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneq/geneq-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneq/geneq-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneq/geneq-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneq/geneq-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneq/geneq-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneq/geneq-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneq/geneq-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneq/geneq-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneq/geneq-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneq/geneq-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneq/geneq-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneq/geneq-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneq/geneq-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneq/geneq-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneq/geneq-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneq/geneq-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneq/geneq-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneq/geneq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2558	100.0	2558	23	ABLO1845
2	1723.4	67.4	6692	23	ABLO1844
3	251.8	9.8	708	23	ABLO2227
4	197	7.7	1584	23	ABLO3499
5	169.8	6.6	2780	23	ABLO2226
6	161.4	6.3	1559	24	ABD24225
7	160	6.3	1303	24	ABLS3644
8	160	6.3	1376	24	ABD24227

C	9	155	6.1	7918	23	ABLO3498	Drosophila melanog
	10	148.2	5.8	829	24	ABK34431	Human CDNA for nov
	11	148.2	5.8	1320	25	AAU54156	Human rhomboid hom
C	12	146	5.7	2589	23	ABLO1854	Drosophila melanog
	13	145.6	5.7	1398	23	ABLO20931	Drosophila melanog
	14	145.6	5.7	1770	23	ABLO3167	Drosophila melanog
	15	108.2	4.2	2858	23	ABLO4764	Drosophila melanog
	16	108.2	4.2	7876	23	ABLO3166	Drosophila melanog
	17	108.2	4.2	9680	23	ABLO20930	Drosophila melanog
	18	107.2	4.2	792	23	ABLO26325	Drosophila melanog
	19	95.2	3.7	1224	24	ABD24226	Human rhomboid rel
	20	95.2	3.7	1600	24	ABO9466	Human coding seque
C	21	94.8	3.7	6668	24	ABLO3697	Human immune syste
C	22	93.6	3.7	14006	24	ABLO33958	Human immune syste
	23	91.6	3.6	1488	25	AAU54157	DNA showing spitz
	24	90.2	3.5	6644	20	AAK33181	Base sequence of t
	25	90.2	3.5	7772	20	AAK33182	Base sequence of t
	26	90.2	3.5	7797	20	AAK33180	Cowpox virus bsr t
	27	90.2	3.5	7996	20	AAK33184	Base sequence of t
C	28	89.6	3.5	3683	25	ABZ10199	Haematopoietic cel
C	29	89	3.5	34769	22	AAU46774	Tumour suppressor
C	30	87.8	3.4	14798	24	ABL33032	Human immune syste
C	31	86.2	3.4	6898	24	ABN80222	Human chemically m
C	32	86	3.4	9539	22	AAU45347	Chemically pretrea
C	33	86	3.4	9539	24	ABK28180	DNA transcription
C	34	85.8	3.4	6292	22	AAU46735	Tumour suppressor
C	35	85.6	3.3	529	25	ABX52180	Bovine EST associa
C	36	85.4	3.3	12337	24	ABL34358	Human immune syste
C	37	85	3.3	15954	24	ABL70468	Chemically treated
C	38	85	3.3	15954	24	ABK31495	Signal transductio
C	39	84.4	3.3	3683	25	ABZ10053	Haematopoietic cel
C	40	84.2	3.3	83391	24	ABO67094	Human angiogenesis
C	41	83.8	3.3	975	24	ABO29508	Oligonucleotide fo
C	42	83.8	3.3	975	24	ABO29509	Human immune syste
C	43	83.4	3.3	8392	24	ABL33490	Human immune syste
C	44	83	3.2	6767	22	AAU46608	Tumour suppressor
C	45	82.6	3.2	6115	24	ABL33801	Human immune syste

ALIGNMENTS

RESULT 1

ID	ABLO1845	standard; cDNA; 2558 BP.
XX	ABLO1845;	
XX	26-MAR-2002 (first entry)	
DE	Drosophila melanogaster expressed polynucleotide SRQ ID NO 17.	
XX	Drosophila; developmental biology; cell signalling; insecticide;	
KW	pharmacological; gene; ss.	
XX	Drosophila melanogaster.	
OS	WO200171042-A2.	
XX	27-SEP-2001.	
PD	23-MAR-2001; 2001WO-US09231.	
PF	23-MAR-2000; 2000US-191637P.	
PR	11-JUL-2000; 2000US-0614150.	
XX	(PEKE) PE CORP NY.	
PA	Venter JC, Adams M, Li PMD, Myers EW;	
XX	WPI; 2001-656660/75.	
DR	P-PSDB; ABB57742.	
XX		

Db	1801	AAATCAAGAGAAAATCAAGAGGAGAAACAAAGACTAATCCCTCGCTGATGATTAAA	1860
Qy	1861	TGATTATTCAGATGTTTTCAATTAATGTTTTGTCTGTTTTCTTAAATTAATGATTTATG	1920
Db	1861	TGATTTTATCCATGTTTTCAATTAATGTTTTGTCTGTTTTCTTAAATTAATGATTTATG	1920
Qy	1921	CCGCAATTAACAGATGGAATCGAATCGAATCAGCAAACTGATCAAAATGTTTTATA	1980
Db	1921	CCGCAATTAACAGATGGAATCGAATCGAATCAGCAAACTGATCAAAATGTTTTATA	1980
Qy	1981	CATCCATTAAGCATATATGTCTCCGAATTTGATGATTTAGTGTATTAATTTATATAATTT	2040
Db	1981	CATCCATTAAGCATATATGTCTCCGAATTTGATGATTTAGTGTATTAATTTATATAATTT	2040
Qy	2041	AGGTTAATCTAGCCCTCCTTAACAAATGTTTCAAATGTATAATCTAATTAATGCGACAC	2100
Db	2041	AGGTTAATCTAGCCCTCCTTAACAAATGTTTCAAATGTATAATCTAATTAATGCGACAC	2100
Qy	2101	TAGTCAAAACAAACAAACAGCAACGCAAAACAAACAAAATATGGAACACACACG	2160
Db	2101	TAGTCAAAACAAACAAACAGCAACGCAAAACAAACAAAATATGGAACACACACG	2160
Qy	2161	AAAGAACATTTCAATTCAGATCAATTAAGCAAACTGAGTAAATTTAATTAATTAATCT	2220
Db	2161	AAAGAACATTTCAATTCAGATCAATTAAGCAAACTGAGTAAATTTAATTAATTAATCT	2220
Qy	2221	AAAGTCACTTAATGCGTTACAAATCGAGCATCGTAATCCCTTACACACACACACACA	2280
Db	2221	AAAGTCACTTAATGCGTTACAAATCGAGCATCGTAATCCCTTACACACACACACACA	2280
Qy	2281	CACACACTCGGAAAGTATCTAATTAATTTAATTTAATGTAAGGCGAGGAGGTTTATTA	2340
Db	2281	CACACACTCGGAAAGTATCTAATTAATTTAATTTAATGTAAGGCGAGGAGGTTTATTA	2340
Qy	2341	ATTGCGTCAATTGAGGGAACATTTATTTATTTATTTAATTTAATTTAGTGAATTCAC	2400
Db	2341	ATTGCGTCAATTGAGGGAACATTTATTTATTTATTTAATTTAATTTAGTGAATTCAC	2400
Qy	2401	ACAAACAAACGACGAAAAACAAACAAACACACAGAGAGAGAGAAACAAACCAATTCAC	2460
Db	2401	ACAAACAAACGACGAAAAACAAACAAACACACAGAGAGAGAGAAACAAACCAATTCAC	2460
Qy	2461	TGTAAAAATATTCCAATTGAAAAATTCACACGAAAAAGCCAAAAAATAAAAATCAAAAC	2520
Db	2461	TGTAAAAATATTCCAATTGAAAAATTCACACGAAAAAGCCAAAAAATAAAAATCAAAAC	2520
Qy	2521	ATTTCAAGATATACACAGTAAATACAAATTCAGACAA	2558
Db	2521	ATTTCAAGATATACACAGTAAATACAAATTCAGACAA	2558

RESULT 2	
ABL01844	
ID	ABL01844 standard; cDNA; 6692 BP.
XX	
AC	ABL01844;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 14.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ss.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.

PR	11-JUL-2000; 2000US-0614150.
XX	(PEKE) PE CORP NY.
PA	Venter JC, Adams M, Li PWD, Myers EW,
PI	WPI: 2001-656860/75.
PX	p-PsDB; ABB57741.
DR	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	Claim 1; SEQ ID NO 14; 21bp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC	sequences (ABL01840-ABL16175) and the encoded proteins
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 6692 BP; 2141 A; 1415 C; 1392 G; 1744 T; 0 other;
Query Match	67.4%; Score 1723.4; DB 23; Length 6692;
Best Local Similarity	93.3%; Pred. No. 0;
Matches 1866; Conservative	0; Mismatches 1; Indels 132; Gaps 1;
QY	TTGAGATTGCATCTTCGCCCTACGACCGCTACACAATGCCCGCAGATTTCGGGCTTAC 751
Db	TTTAGATTGCAATCTTGCTGCTTAAGACCGGTACAAATGCCCGCAGATTTCGGGCTTAC 3753
QY	752 CCGTTCGGATTCCGTCGGATTCCGTCGTCTATCGGCGGACCGGCGTCTGCAAGTGT 811
Db	CCGTTCCGATTCCGTCGGATTCCGTCGTCTATCGGCGGACCGGCGTCTGCAAGTGT 3813
QY	3754 CGGTCGGATTCCGTCGGATTCCGTCGTCTATCGGCGGACCGGCGTCTGCAAGTGT 812
Db	GCGGCTCTTGAATCATCTTCCTGACCGCAACCTGGTTCACCTGGGCTTCAATATCG 871
QY	812 GCGGCTCTTGAATCATCTTCCTGACCGCAACCTGGTTCACCTGGGCTTCAATATCG 871
Db	GCGGCTCTTGAATCATCTTCCTGACCGCAACCTGGTTCACCTGGGCTTCAATATCG 3873
QY	872 TCATCCAGCTGTTTTTGGGCAATCCCCTGAGGAGTACGACGCGCACCGCAGATTCGGCG 931
Db	TGATCCAGCTGTTTTTGGGCAATCCCCTGAGGAGTACGACGCGCACCGCAGATTCGGCG 3933
QY	932 TGATTCATCATGCGGGGCGTTTTTGC CGGATCCCTGGGACCAAGTGTCTGACTCGGAGG 991
Db	TGATTCATCATGCGGGGCGTTTTTGC CGGATCCCTGGGACCAAGTGTCTGACTCGGAGG 3934
QY	992 TCTTTCCTGGTGGGCGCGCAGGCGGTGCGTCTATATGCCCTGTGGCGGACATCTGGCCACA 1051
Db	TCTTTCCTGGTGGGCGCGCAGGCGGTGCGTCTATATGCCCTGTGGCGGACATCTGGCCACA 4053
QY	1052 TTACACTGAATATGCGACATGAGAAGCGCATCCAGCAACTCGAATCCGTTGTCACT 1111
Db	TTACACTGAATATGCGACATGAGAAGCGCATCCAGCAACTCGAATCCGTTGTCACT 4113
QY	1112 TT----- 1113
Db	TTGGTGAATTTCCCAACACCTTAATGACTAGATCTATAGAACCTATCTACTACTAA 4173
QY	1114 ----- 1113
Db	ACAAGATTGTGCAATCTTTTCTCTATGTTAACTTAACATACTGAATCCCTTCTG 4233
QY	1114 -----GTCTCCGAGATCGGAGTATGCTCTTACACCATACTTCTGANT 1159
Db	TTCTTCATTGCAATCTCTCGGATCTGGGCTATGCTCTTACACCATACTTCTGANT 4293

QY 1160 GAAGCGCTTCGCCAAGGATGCCAGGTGTCTAATTCAGTCCCACTGACGGAGCCCTGG 1219
DB 4294 GAAGCGCTTCGCCAAGGATGCCAGGTGTCTAATTCAGTCCCACTGACGGAGCCCTGG 4353
QY 1220 CAGGACTAACGATCGGCTTCTGGTGTCTAAGAACTTCGGTCAACCGAGTACGAGCAGC 1279
DB 4354 CAGGACTAACGATCGGCTTCTGGTGTCTAAGAACTTCGGTCAACCGAGTACGAGCAGC 4413
QY 1280 TCATCTGGTGGCTAGAGGCTGGGGGTCTACGTGCTTCAACGGCTTCGCGCATGGTTTCA 1339
DB 4414 TCATCTGGTGGCTAGAGGCTGGGGGTCTACGTGCTTCAACGGCTTCGCGCATGGTTTCA 4473
QY 1340 ACCTGATCAACACGGTGACCGCTCAGCTGATGAGAGACGAGGTGAGTATTAACCGAGC 1399
DB 4474 ACCTGATCAACACGGTGACCGCTCAGCTGATGAGAGACGAGGTGAGTATTAACCGAGC 4533
QY 1400 ATCTGTTGACGACCTGGGAGTGTCTTAAGTTGAGGTTGGAGTGTCTGACATGCTGCG 1459
DB 4534 ATCTGTTGACGACCTGGGAGTGTCTTAAGTTGAGGTTGGAGTGTCTGACATGCTGCG 4593
QY 1460 AGGGAATTCGGAATCTGCTTACGCTTCAGAGAGATCGAGAGACAGAGATGGTGGAAAA 1519
DB 4594 AGGGAATTCGGAATCTGCTTACGCTTCAGAGAGATCGAGAGACAGAGATGGTGGAAAA 4653
QY 1520 GAAAAGTTCACTCAACGATTTAGTTCAAAATTAATTCGATTTCTTTGGCTTTTGGCTTT 1579
DB 4654 GAAAAGTTCACTCAACGATTTAGTTCAAAATTAATTCGATTTCTTTGGCTTTTGGCTTT 4713
QY 1580 TCGTTAGCATTTATCTGTTATCTGTTACGTTTGCAGTTAAACGTTTCCGAAACAT 1639
DB 4714 TCGTTAGCATTTATCTGTTATCTGTTACGTTTGCAGTTAAACGTTTCCGAAACAT 4773
QY 1640 AGTACCAAACTCAAAAACCAAAACCAATCAAGAGAAATCACTGACCAAAAAGAAA 1699
DB 4774 AGTACCAAACTCAAAAACCAAAACCAATCAAGAGAAATCACTGACCAAAAAGAAA 4833
QY 1700 GTGCGAGAGATGAGAGAACATTAACCGAAGCCGAAACGTGTAAACAAATGTTGTAG 1759
DB 4834 GTGCGAGAGATGAGAGAACATTAACCGAAGCCGAAACGTGTAAACAAATGTTGTAG 4893
QY 1760 AACCAAACTGAATTTATTTTCCGCTGTAAAAACCAAGTAAAAATCAAGAGAAATCAA 1819
DB 4894 AACCAAACTGAATTTATTTTCCGCTGTAAAAACCAAGTAAAAATCAAGAGAAATCAA 4953
QY 1820 AGAGAGAAACAGAACTTAATCGCTCTCGCTAGATTAATTAATGATTAATCAAGTTTCA 1879
DB 4954 AGAGAGAAACAGAACTTAATCGCTCTCGCTAGATTAATTAATGATTAATCAAGTTTCA 5013
QY 1880 ATTATGTTTGTCTGTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1939
DB 5014 ATTATGTTTGTCTGTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5073
QY 1940 ATCGAATCGAAGCATCGAACTGTATCAAAATGTTTATATCATCATTAACATATGTG 1999
DB 5074 ATCGAATCGAAGCATCGAACTGTATCAAAATGTTTATATATCATCATTAACATATGTG 5133
QY 2000 CTCGGATTTGAGATTTAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2059
DB 5134 CTCGGATTTGAGATTTAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 5193
QY 2060 AACAAATTTGTTCAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2119
DB 5194 AACAAATTTGTTCAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5253
QY 2120 GCAACAGCAAAAACCAAAAACCAAAAACCAAAAACCAAGCAAAACATTCATTTAG 2179
DB 5254 GCAACAGCAAAAACCAAAAACCAAAAACCAAAAACCAAGCAAAACATTCATTTAG 5313
QY 2180 ATCAATTAAGCAATTCAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2239
DB 5314 ATCAATTAAGCAATTCAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5373
QY 2240 CAAAATCGAGATGTAATCCCTTACACACACACACACACACACACACTCGAAAGTATTA 2299

DB 5374 CAAAATCGAGATGTAATCCCTTACACACACACACACACACACACACTCGAAAGTATTA 5433
QY 2300 CTAAATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2359
DB 5434 CTAAATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5493
QY 2360 TATTTATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2419
DB 5494 TATTTATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5553
QY 2420 CAACAAACACAG 2479
DB 5554 CAACAAACACAG 5613
QY 2480 AAAATACACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 2539
DB 5614 AAAATACACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 5673
QY 2540 AATAACAAATTAACAGCAA 2558
DB 5674 AATAACAAATTAACAGCAA 5692

RESULT 3
ID ABL02227
ABL02227 standard; cDNA; 708 BP.
XX
AC ABL02227;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 1163.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN W0200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
PI Venter JC, Adams M, Li PWD, Myers EM;
XX
DR WPI, 2001-656860/75.
XX
PT P-PSDB; ABB58124.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 1163; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB5737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

QY 1221 AGAGTAACGATCGCTTCTGTGCTAAGAATTGCGTCAACGAGTACGAGAGCT 1280
DB 789 TGGCTTAATATCGGGCTGTGTCTCAAGAGCTTTGAGCAGAGCTCCAGAGAGCT 848
QY 1281 CATCTGGTGGCTAGCGCTTGGGCGCTCTAGCTGTCTTACCGCTTGGCCATGTTTCA 1340
DB 849 GCTCTGTGTGATCGCTTGGGAAAGTACTGCGCTGTGTCTTGTGCAATGCTTCA 908
QY 1341 CCTGATCA 1349
DB 909 CATCATGA 917
RESULT 5
ABL02226
ID ABL02226 standard; cDNA; 2780 BP.
XX ABL02226;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 1160.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PP
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX P-PsDB; ABB58123.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PS
XX Claim 1; SEQ ID NO 1160; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 2780 BP; 744 A; 637 C; 645 G; 754 T; 0 other;
SQ
Query Match 6.6%; Score 169.8; DB 23; Length 2780;
Best Local Similarity 56.2%; Pred. No. 1.8e-28;
Matches 413; Conservative 0; Mismatches 247; Indels 75; Gaps 2;
QY 763 CGCTGGATTCGGTCTGTCTATCGCCGAGCCGGCTGTGCAAGTGTGGCGCTTTCTT 822
DB 1040 CCAGAGGATTCCTTCTGTGTATCGACCCGACGAGCGCTGTGAGATTCCTT 1099
QY 823 AGTCAATGTTCTCGACGCACTGGTCCACTGGGCTTCAATATGTCATCACTG 882

DB 1100 ACCTATGCCCTCTCCAGCCAGTTGGCTTCAATCTGGTTACAAATGTCTCACTAGCTC 1159
QY 883 TTTTGGCATTCCTCCAGAGGTGATGCAAGGCAAGGATGCGCTGATCTACATG 942
DB 1160 CTGTTCCGATTCCTCAATTGAGCTGTGATGATGCTGAGAACCGGATCATTTATATG 1219
QY 943 GCGGCGTTTTTTCGCGATCCCTGGGCAACAGTGTCTGTGACTCGAGAGTCTTCTGCTG 1002
DB 1220 GCGGAGTTTTTGTGCTGCTCTTTGGGAACCTTCCGTGTGAGTCTCGAAGTTTTTCTGCTG 1279
QY 1003 GCGCCAGCGGTGGGTCTATATCCCTGTGGCCGACATTTGGCCAACTTACCTGAGC 1062
DB 1280 GGAGCCAGTGAAGAGTTTACGCTTGTGGCCCAACTGCGCATTTTGTCTCAAC 1339
QY 1063 TATGGGCATGAAAGAGCATCCAGCACTCGATCCGTGTGATCTT----- 1112
DB 1340 TTTGTCAATATGACACGAGGTAATTACGTTAATGCTGTGATTTTGTCCGTAAGTGA 1399
QY 1113 ----- 1112
DB 1400 ACTGATATATTCTAATATGAATTTGCTGTAGATCATAAACATTTCTATCTTT 1459
QY 1113 --TGTCTCTGGATCTGGGCTATGCTCTTCAACCCAACTACTGATGAAAGCGCTTC 1170
DB 1460 ACAGTTTCTGCGATTTGGGATGATGCTCTTATTAAGAACTGCAATGATCAGCTG 1519
QY 1171 GCCAAGGATCCCAAGTGTGTATGATTCGCACTGACGAGGAGCCCTGACAGACTAAG 1230
DB 1520 CAGACAGACCTCTGTCTCTATATGCTCAATGACCGGTGTCTGCGGATTAAT 1579
QY 1231 ATCGCTTCTGTGCTAAGAATTGCTGCAACGAGTACAGAGCTCATCTGTGG 1290
DB 1580 GTGGGTCTCTCTCTTCCGGAACGATGATGGGCTGCAACAGCCCACTCCGCTGG 1639
QY 1291 CTAGCGTTGGGCGTCTACTGTGCTTCAACGCTTTCGATGTTTCAACTGATCAAC 1350
DB 1640 TTGGCCCTGGGCGTATGTGATTAATCAGCGCTTGGATGCGCTTCAACTGTGAT 1699
QY 1351 ACGGTGACCGCTCAGCTGATG---GAGGAGAGGAGTGAATTCAGCAGATCTGTTG 1407
DB 1700 ACGGTCAAGCTCACTTCTGTGCGGAGAGAGGAGGAGGATGATCAGGACATCTGATG 1759
QY 1408 CACGACCTGGAGTG 1422
DB 1760 AATGACTCGAATG 1774
RESULT 6
AAD24225
ID AAD24225 standard; cDNA; 1559 BP.
XX
XX AAD24225;
AC
XX 17-MAY-2002 (first entry)
DT
XX Human rhomboid related protein, RRP1 cDNA.
DE
XX Human; rhomboid related protein; RRP1; transmembrane protein;
KM epidermal growth factor receptor; EGFR signalling pathway;
KM tissue growth factor alpha; TGFalpha; epidermal growth factor; EGF;
KM amphiregulin; tumour; p21; cancer; gene therapy; ss.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 28..1344
FT CDS /tag= a
FT /product= "rhomboid related protein, RRP1"
PN WO200205843-A2.
XX 24-JAN-2002.


```

QY 711 CTACGACCGCTACACAAATGCCGCCGAAATTCGGGCTACCCGTTCCGATTCCGTCGGA 770
DB 388 CTACATAGGGGTGTATCGATAGTCAATTTGTATCTGACAGATTAATCATCAGTACTTGAA 447
QY 771 TTGGGTGCTGTGATCGGCGGACGGCGTCTGACAGTGTGGCGCTTTTGTAGTCAAT 830
DB 448 GAACTCCCTGTTTACCAACCAAGCTGACAGCAGAGTTGGGCTTACCTGACATACAT 507
QY 831 GTTCTGACGCGCAATGTTGCTTCCAGCTGAGCTTCAATATGTCATCCAGCTGTTTTTGG 890
DB 508 CTTCATGATGACGAGGATAGAACACCTGGGACTAATGTGTGTGTGAGCTGCTGGTGGG 567
QY 891 CATTCCTGAGAGTGTGACAGGACGCGCAGAGATGCGGTGATCTTACATGCGGCGGT 950
DB 568 GGTGCCCCCTGAGATGTGATGAGGACCCGCAATGGGCTTGTCTTACCTGCGCGGTGT 627
QY 951 TTTTGCCGATTCCTGCGGACACAGTGTCTGCTGAGCTGCGAGAGTCTTCTGTGTGGCGCAG 1010
DB 628 TGTGGCAGGGTCTTGGCAGTGTCTGTGCTGACATACCGCTCCAGTGTGGGCTTTC 687
QY 1011 CGGTGCGCTATGACCTGTTGGCCGACATCTGGCCAACTTACACTGAACTATGCGCA 1070
DB 688 TGAAGGGGTGTATGCTCTGCTCTGCTGCTGCGCAATCTGCAATGTCTATGAACTGTGAG 747
QY 1071 CATGAAAGCGCATTCACGCAACTCGGATCGTTGTATCTTGTCTCTGCGATCTGGG 1130
DB 748 CATGAAGTGCAGATTCAGCTGCTGCGAGATGCGTGTGGCCCTTATCTGTATGAGCATGGA 807
QY 1131 CTATGCTCTTACACC---CAATATCTGATGAGAGGCGCTTGGCCAAAGGTCCCAAGT 1187
DB 808 GTTTGGCGGGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 867
QY 1188 GTGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1247
DB 868 AACCTTTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 927
QY 1248 AAGAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1307
DB 928 GAGGAATCTGAGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 987
QY 1308 CTGTGCTTACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343
DB 988 CACCGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023

RESULT 8
AAD24227 standard; cDNA; 1376 BP.
AC AAD24227;
DT 17-MAY-2002 (first entry)
XX Human rhomboid related protein, RRP3 cDNA.
XX Human rhomboid related protein, RRP3; transmembrane protein;
XX epidermal growth factor receptor; EGFR signalling pathway;
XX tissue growth factor alpha; TGFalpha; epidermal growth factor; EGF;
XX amphiregulin; tumour; p21; cancer; gene therapy; ss.
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
XX FH 248..1336
XX FT /*tag= a
XX PD /product= "Rhomboid related protein, RRP3"
XX PF 18-JUL-2001; 2001MO-US22555.
XX 24-JAN-2002.
XX WO200205843-A2.
XX 18-JUL-2001; 2001MO-US22555.

```

```

PR 19-JUL-2000; 2000US-219289P.
PR 21-MAR-2001; 2001US-277471P.
PR 21-MAR-2001; 2001US-277471P.
PR 12-JUL-2001; 2001US-304863P.
XX (EXEL-) EXELIXIS INC.
XX Friedman L, Larson JS, Belvin M, Chen C, Robertson SA, Lioubin M;
XX Shi W, Chan J, Li D;
XX WPI; 2002-179749/23.
XX P-PSDB; AAE14571.
XX Human rhomboid related proteins involved in epidermal growth factor
XX receptor signalling pathway, useful for screening compounds modulating
XX interaction between the protein and binding target, and for inhibiting
XX growth of tumors
XX Claim 22; Page 34; 36pp; English.
XX The invention relates to human rhomboid related proteins (RRPs) which
XX are transmembrane proteins involved in epidermal growth factor receptor
XX (EGFR) signalling pathway. The polypeptide of the invention is useful for
XX screening agents that modulate the interaction of RRP with an RRP binding
XX target (e.g. tissue growth factor (TGF)alpha, epidermal growth factor
XX (EGF), and amphiregulin). The RRP modulating agent is an antibody or a
XX small organic molecule which can be used to specifically inhibit growth
XX of tumor cells that overexpress an RRP protein. The RRP is important
XX drug targets for the treatment of pathologies associated with a defective
XX p21 signalling pathway, e.g. cancer. RRP polynucleotide is useful for
XX diagnosing a tumor cell, in gene therapy, in genetic studies to
XX elucidate the EGFR pathway, and in animal models of disease and disorders
XX implicating defective EGFR pathway function. The polynucleotide is also
XX useful as translatable transcripts, hybridisation probes and PCR primers.
XX The hybridisation probes are used for identifying wild-type and mutant
XX alleles in clinical and laboratory samples. Mutant alleles are
XX used to generate allele-specific oligonucleotide (ASO) probes for
XX high-throughput clinical diagnoses. The present sequence is human
XX RRP3 cDNA.
XX Sequence 1376 BP; 276 A; 378 C; 415 G; 307 T; 0 other;
XX Query Match 6.3%; Score 160; DB 24; Length 1376;
XX Best Local Similarity 52.9%; Pred. No. 2.3e-26;
XX Matches 368; Conservative 0; Mismatches 325; Indels 3; Gaps 1;
QY 651 CCGCCAGCACTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 710
DB 595 CTGCTGCCCCCACCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 654
QY 711 CTACGACCGCTACACAAATGCCGCCGAAATTCGGGCTACCCGTTCCGATTCCGTCGGA 770
DB 655 CTACATAGGGGTGTATCGATAGTCAATTTGTATCTGACAGATTAATCATCAGTACTTGAA 714
QY 771 TTGGGTGCTGTGATCGGCGGACGGCGTCTGACAGTGTGGCGCTTTTGTAGTCAAT 830
DB 715 GAACTCCCTGTTTACCAACCAAGCTGACAGCAGAGTTGGGCTTACCTGACATACAT 774
QY 831 GTTCTGACGCGCAATGTTGCTTCCAGCTGAGCTTCAATATGTCATCCAGCTGTTTTTGG 890
DB 891 CATTCCTGAGAGTGTGACAGGACGCGCAGAGATGCGGTGATCTTACATGCGGCGGT 950
QY 951 TTTTGCCGATTCCTGCGGACACAGTGTCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1010
DB 835 GGTGCCCCCTGAGATGTGATGAGAGCCACCGAATGGGCTTGTCTTACGTGCGCGGTGT 894
QY 895 TGTGGCAGGGTCTTGGCAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
QY 1011 CGGTGCGCTATGACCTGTTGGCCGACATCTGGCCAACTTACACTGAACTATGCGCA 1070
DB 955 TGAAGGGGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014

```


XX Claim 1, Page 125-126, 339pp, English.
 PS
 CC The invention relates to 625 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins, their complements and sequences that hybridise to them.
 CC Also included are a vector comprising the polynucleotide, a host cell
 CC transformed with the vector, the proteins encoded by the
 CC polynucleotides, antibodies that bind to the proteins and identification
 CC of modulators of the proteins or the expression of the polynucleotide.
 CC The polynucleotides can be used as probes for the identification
 CC and isolation of full length cDNA and genomic DNA. The polynucleotides
 CC and proteins can also be used as nutritional supplements. The protein
 CC is useful in the treatment of various immune deficiencies and disorders
 CC such as viral infections, bacterial infections, fungal infections,
 CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
 CC autoimmune thyroiditis and diabetes) and allergic reactions and
 CC conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment
 CC of burns, incisions and ulcers. The proteins are also useful for
 CC regulating haematopoiesis, for treating myeloid or lymphoid cell
 CC deficiencies. The present sequence is one of the 625 cDNA sequences
 CC encoding a secreted protein.
 XX
 XX Sequence 829 BP; 144 A; 230 C; 259 G; 196 T; 0 other;

Query Match 5.8%; Score 148.2; DB 24; Length 829;
 Best Local Similarity 55.0%; Pred. No. 9e-24;
 Matches 313; Conservative 0; Mismatches 253; Indels 3; Gaps 1;

QY 778 CTGGTCTATCGGCGGACCGGCGCTGACAGTGTGGGCTTTTACTACATGTTCCG 837
 DB 26 CTGGTTTACACACCGACGCTGCGAGCACAGGTTTGGCTTACCTGACATCATCTTATG 85
 QY 838 CAGGCCAATGTTTCCATCTGGGCTTCAATATGCTCATTCAGCTGTTTGGCATCC 897
 DB 86 CATGACAGGATAGAACCTGGGACTCAATGTGTGTGACGCTGTGGGGTCC 145
 QY 898 CTGAGAGTGTATGACGCGACCGGACGAGTGGCGTGTATCATGAGCGGCGTTTGGCC 957
 DB 146 CTGAGAGTGTGTATGAGGACCGGAAATGGGCTGTATACGTCGCGGATGTGTGCA 205
 QY 958 GGAATCCTGGGACCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017
 DB 206 GGGTCTTGGGCAAGT 265
 QY 1018 GTCTATGCTGTTGGGCGGACATCTGCGCAATATACATGACATGACATGAG 1077
 DB 266 GTGTATGCTGCTGCTGTGCGCATCTGCGCAATGTGTGTGTGTGTGTGTGTGTGT 325
 QY 1078 AGCGCATGCAAGCAATCGGATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1137
 DB 326 TGGCAGTTTCAAGCTGTGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 385
 QY 1138 C---TTCACACCAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1194
 DB 386 CGGCGCTGT 445
 QY 1195 ATTGCCACCTGACGAGGAGCGCTGCGAGCATACATGAGCTTGTGTGTGTGTGTGT 1254
 DB 446 GTGGCGCATTTGGT 505
 QY 1255 TTGGTATCAGAGAGTATGAGACAGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1314
 DB 506 TAGAGAGAGAGGCTCCAGACAGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 565
 QY 1315 TTACCGTCTTGGCATGCTTTTCAACT 1343
 DB 566 TTGCTGCTGT 594

RESULT 11
 ID AAL54156
 AC AAL54156; standard; DNA; 1320 BP.
 DT 28-MAR-2003 (first entry)
 DE Human rhomboid homologue RHBDL-3 DNA.
 XX
 XX Cytostatic; antiarteriosclerotic; antipsoriatic; vulnerary; nootropic;
 XX neuroprotective; protease modulator; modulator; rhomboid; EGF receptor;
 XX proteolytic cleavage; ErbB; cancer; coronary atherosclerosis; psoriasis;
 XX wound healing; infant prematurity; peripheral nerve injury; neuropathy;
 XX gene therapy; human; chr 11; ds.
 OS Homo sapiens.
 XX
 XX MO200293177-A2.
 XX
 XX 21-NOV-2002.
 XX
 XX 13-MAY-2002; 2002MO-GB02234.
 XX
 XX 11-MAY-2001; 2001GB-0011574.
 XX
 XX 27-SEP-2001; 2001GB-0023261.
 XX
 XX (MED1-) MEDICAL RES COUNCIL.
 XX
 XX Freeman M;
 XX
 XX WPI; 2003-129326/12.
 XX
 PT Identifying a modulator of a rhomboid polypeptide, useful for treating
 PT cancer, coronary atherosclerosis, or psoriasis, comprises contacting a
 PT rhomboid polypeptide and a polypeptide substrate in the presence of a
 PT test compound -
 XX
 PS Claim 18; Fig 7; 110pp; English.

XX The invention relates to a novel assay for identifying a modulator of a
 XX rhomboid polypeptide comprising contacting a rhomboid polypeptide and a
 XX polypeptide substrate in the presence of a test compound, and determining
 XX the proteolytic cleavage of the polypeptide substrate. The polypeptide,
 XX polypeptide fragment, and modulator are useful in the manufacture of a
 XX composition for treating a condition associated with aberrant ErbB or EGF
 XX receptor activity or a pathogen infection, e.g. cancer, coronary
 XX atherosclerosis, psoriasis, wound healing, infant prematurity, or a
 XX peripheral nerve injury/neuropathy. The polynucleotides of the invention
 XX can be used in gene therapy. This polynucleotide sequence represents the
 XX DNA of a human rhomboid homologue RHBDL-3 relating to the novel assay of
 XX the invention.

QY Sequence 1320 BP; 256 A; 383 C; 412 G; 269 T; 0 other;

Query Match 5.8%; Score 148.2; DB 25; Length 1320;
 Best Local Similarity 55.0%; Pred. No. 1.1e-23;
 Matches 313; Conservative 0; Mismatches 253; Indels 3; Gaps 1;

QY 778 CTGGTCTATCGGCGGACCGGCGCTGACAGTGTGGGCTTTTACTACATGTTCCG 837
 DB 709 CTGGTTTACACACCGACGCTGCGAGCACAGGTTTGGCGTACCTGACATCATCTTATG 768
 QY 838 CAGGCCAATGTTTCCATCTGGGCTTCAATATGCTATCCAGCTGTTTGGCATTTCC 897
 DB 769 CATGACAGGATAGAACCTGGGACTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 828
 QY 898 CTGAGAGTGTATGACGCGACGCGGAGTGGCGTGTATCATGAGCGGCGGCTTTTGGCC 957
 DB 829 CTGAGAGTGTGTATGAGGACCGGAAATGGGCTGTCTACGTTGGCGGATGTGTGCA 888
 QY 958 GGAATCCTGGGACCAAGT 1017

DB 889 GGTTCCTTGGCACTGCTGTGGCTGACATGACCGCTCCAGTCCGTGGCTCTTCTGAGGG 948
1018 GTCTATGACCTGTGGCCGCACTGACCACTTACCTGATGAGCATGAGG 1077
949 GTGTATGCTCTCTCTCTGCCCCATCTGGCCACATTTGATGACTGCTGAGCATGAGG 1008
1078 AGCGATCCACGCACTCGGATCCGTTCATCTTGTCTCCGATCTGGCATGTGT 1137
1009 TCGCATTTCAAGCTGTGGCATGAGCTGTGGCCCTTATCTGTGTGAGCATGAGATTGGG 1068
1138 C---TTTACACCAATCTTCTGATGAGAGGCTTTGGCCAAAGGTCTCCAGGTGTGAC 1194
1069 CGGCGCTGTGGCTCCGCTTCCACCCGCTGATCCCGGCTTACCCCAACCACTTT 1128
1195 ATTGACCACTGACGGGAGCCCTTGACGACATGAGCTGCTTGTGGCTTAAGAAC 1254
1129 GTGGCGCATTTGGTGGCTGGCGCTGGCGGATCACTTGGCGGTGGTCTGAGGAGAC 1188
1255 TTGCGTACCGAGATGACGACGCTCATCTGTGGCTGACGCTTGGGCGCTTACTGTGCC 1314
1189 TACGACCAAGGCTCCAGGACCACTGACTGTGGTGAATTTTGTGCGCATGTACACCGTC 1248
1315 TTACCGCTCTTGGCCATCGTTTCAACT 1343
1249 TTGCTGTGTGCTGCTGCTTCTGGAACAT 1277

RESULT 12

ABL01854/c
ID ABL01854 standard; cDNA; 2589 BP.

AC ABL01854;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 44.

KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical; gene; ss.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PERK) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR MPI: 2001-656860/75.

DR P-PSDB; ABB57751.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

PS Claim 1; SEQ ID NO 44; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL01854-ABL16175), expressed DNA
sequences (ABL01854-ABL16175) and the encoded proteins
(ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX Sequence 2589 BP; 772 A; 595 C; 558 G; 664 T; 0 other;

Query Match 5.7%; Score 146; DB 23; Length 2589;
Best Local Similarity 65.2%; Pred. No. 4,8e-23;
Matches 215; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 720 CTACCAATATCCCGCCCAATTTGGGCTACCCGTTCCGATTCGCGGATTCGTCT 779
DB 334 CCACCTAGATGTCACGGGTGAGCGCGCCGCAAGGAGACCATACCTCGACATGTT 275
QY 780 GGTCTATGCGCGGACCGGCGCTGTGAGGTGAGCGCTTCTTACGATCATGTTCCGCA 839
DB 274 CATCTACGATCCCGATTAAGCTGACAGATCTGGGCTTCTCTTACATGTAATCTCA 215
QY 840 CGCCAACTGTTTCACTGAGGCTTCATATCTCATCCAGCTGTTTGTGGCATTTCCCT 899
DB 214 TCCCGGATGCTGACATCTCGTTTAAAGTGCAGTTGAGCTGTTTGGGCTGCACT 155
QY 900 GGAAGTATGACGACGCGACGCGGATGCGGTGATCTAATGCGGCGCTTTTCCGG 959
DB 154 GGAAGATGTTTCAATGTTCCACGAGGATGCTGACATTAATCTTCCGGGCTGCTGGCG 95
QY 960 ATCCCTGAGGACAGATGCTGACCTGCGAGGCTTCTCGTGGGCGGCGGCTGCGCT 1019
DB 94 AAGCTTGGGACATTAAGATTTTGAACCGGATGTGTTCTTGTGGGCGGCGGATGAGCT 35
QY 1020 CTATGCGCTGTGGCGGACATCTGGCCAA 1049
DB 34 ATACGCTTGTGACGCGGCCCACTGGCCAA 5

RESULT 13

ABL20931
ID ABL20931 standard; DNA; 1398 BP.

AC ABL20931;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 14266.

KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical; gene; ds.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PERK) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR MPI: 2001-656860/75.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

PS Claim 1; SEQ ID NO 14266; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1398 BP; 315 A; 382 C; 363 G; 338 T; 0 other;

Query Match 5.7%; Score 145.6; DB 23; Length 1398;
Best Local Similarity 54.8%; Pred. No. 4.5e-23;
Matches 320; Conservative 0; Mismatches 249; Indels 15; Gaps 1;

QY 763 CCGTCGGATTGCGTGTCTATCGCCGAGCCGGGTCTGCAAGTGTGCGCTTTCTT 822
DB 793 CCGCGCGCCACCTGTTTCAATCCATACAGGCGTATGAGGATGCGCTTTGTT 852
QY 823 AGCTACATGTTCCGACGCGCAACTGGTTCACCTGGGCTTCAATATCGTATCCAGCTG 882
DB 853 AGCTACATGTTCCGACGCGCAACTGGTTCATGATGATGATGATGATGATGATGATG 912
QY 883 TTTTGGGCAATCCCTGAGAGTGTATGACAGCGCAAGATCGGCGTATCTACATG 942
DB 913 TTTCTGGGCAATCCCTGAGAGTGTATGACAGCGCAAGATCGGCGTATCTACATG 972
QY 943 GCGGCGCTTTTGGCGGATCCCTGGGAGCAAGTGTCTGCAATCGGAGTCTTCTG 1002
DB 973 GCTGGCGTGTAGCCGCTTCCATGCGGCACTCGCTAACAGTCTCGCATCTTCTAGCG 1032
QY 1003 GCGGCGAGCGTGGCGCTTATGCGCTGTTGGCGGCAATCGGCAATTAACAGTGAAC 1062
DB 1033 GCGGCGATCGGTGAGTATGACGATGATGATGATGATGATGATGATGATGATGATG 1092
QY 1063 TATGCGCAATGAGAGCGGATCCAGCAACTCGGATCGGATCGGATCGGATCGGATCG 1122
DB 1093 TACTCGGAATGAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATG 1152
QY 1123 GATCTGGGATGCTCTTACACCAATPACTTGATGAAAGCGCTTGGCCAGGGTCC 1182
DB 1153 GACTTGGGATGCTCTTACACCAATPACTTGATGAAAGCGCTTGGCCAGGGTCC 1197
QY 1183 CAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1242
DB 1198 CAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1257
QY 1243 GTGCTAAGAACTTCGCTACCGAGAGTACGAGAGTACGAGAGTACGAGAGTACGAGAGT 1302
DB 1258 GTGCTGCGCAATTTGAGAGTGTGGGCGGATGGGAGCGCATTTCTGTGGTGTGGCGTATC 1317
QY 1303 GTCTACTGTGCTTACCGCTTCTTCCCATGCTTTTCAACCTGAT 1346
DB 1318 GTTACTTGTGCGCTTACGAGCACTGCGATCATCATACAGGTGTT 1361

RESULT 14

AB103167
ID AB103167 standard; cDNA; 1770 BP.

XX AC ABL03167;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 3983.

XX KM Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical; gene; 88.

XX OS Drosophila melanogaster.

XX PN MO200171042-A2.

XX 27-SEP-2001.
PD 23-MAR-2001; 2001WO-US09231.
XX PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX PR
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EM;
PI
XX WPI; 2001-656860/75.
DR P-PSDB; ABB59064.
XX
XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PS
XX Claim 1; SEQ ID NO 3983; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1770 BP; 416 A; 493 C; 428 G; 433 T; 0 other;

Query Match 5.7%; Score 145.6; DB 23; Length 1770;
Best Local Similarity 54.8%; Pred. No. 5e-23;
Matches 320; Conservative 0; Mismatches 249; Indels 15; Gaps 1;

QY 763 CCGTCGGATTGCGTGTCTATCGCCGAGCCGGGTCTGCAAGTGTGCGCTTTCTT 822
DB 867 CCGCGCGCCACCTGTTTCAATCCATACAGGCGTATGAGGATGCGCTTTGTT 926
QY 823 AGCTACATGTTCCGACGCGCAACTGGTTCACCTGGGCTTCAATATCGTATCCAGCTG 882
DB 927 AGCTACATGTTCCGACGCGCAACTGGTTCATGATGATGATGATGATGATGATGATG 986
QY 883 TTTTGGGCAATCCCTGAGAGTGTATGACAGCGCAAGATCGGCGTATCTACATG 942
DB 987 TTTCTGGGCAATCCCTGAGAGTGTATGACAGCGCAAGATCGGCGTATCTACATG 1046
QY 943 GCGGCGCTTTTGGCGGATCCCTGGGAGCAAGTGTCTGATCGGAGTCTTCTG 1002
DB 1047 GCTGGGCTGTATCCGCTTCCATGAGGCACTCGCTAACAGTCTTCTGATCTTCTAGCG 1106
QY 1003 GCGGCGAGCGTGGCGCTTATGCGCTGTTGGCGCAATCTGCGCAATTAACAGTGAAC 1062
DB 1107 GCGGCGATCGGTGAGTATGACGATGATGATGATGATGATGATGATGATGATGATG 1166
QY 1063 TATGCGCAATGAGAGCGGATCCAGCAACTCGGATCGGATCGGATCGGATCGGATCG 1122
DB 1167 TACTCGGAATGAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATG 1226
QY 1123 GATCTGGGATGCTCTTACACCAATPACTTGATGAAAGCGCTTGGCCAGGGTCC 1182
DB 1227 GACTTGGGATGCTCTTACACCAATPACTTGATGAAAGCGCTTGGCCAGGGTCC 1271
QY 1183 CAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1242
DB 1272 CAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1331
QY 1243 GTGCTAAGAACTTCGCTACCGAGAGTACGAGAGTACGAGAGTACGAGAGTACGAGAGT 1302

Db 1332 GTGCTGCGCAATTGGAGGTGCGGCGATGAGGACGCAATCTCTGTGGTCCGCTCATTG 1391
QY 1303 GTCTACTGTGCTTCAACCGCTTCCGCATCGTTTCAACCTGAT 1346
Db 1392 GTTTACTTTGCTCTCATGACCACTGGCATTCATATACAGTGTT 1435

RESULT 15
ABL04764
ID ABL04764 standard; cDNA; 2858 BP.
XX ABL04764;
AC ABL04764;
XX 26-MAR-2002 (first entry)
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 8774.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX OS
XX MO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EM;
XX WPI; 2001-656860/75.
XX DR P-PSDB; ABB60661.
XX CC

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX

PS Claim 1; SEQ ID NO 8774; 21bp + Sequence Listing; English.
XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 2858 BP; 675 A; 671 C; 668 G; 844 T; 0 other;
XX

Query Match 4.2%; Score 108.2; DB 23; Length 2858;
Best Local Similarity 60.3%; Pred. No. 2.1e-14;
Matches 179; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
XX

QY 763 CGGTGGATTGCGTGTGCTATGCGCGGACCGGCGTCTGAGGTGCGGCTTCTT 822
Db 2446 CCAGCGGCCACCTGTTCATCTACATCCGTACACGCGCTATGAGGATGGCGCTTGT 2505

QY 823 AGCTACATGTTCTGCGACCGCACTGTTTCACTGGGCTTCAATTCCTCATCCAGCTG 882
Db 2506 AGCTACATGTTCTGCGACCGCACTGTTTCACTGGGCTTCAATTCCTCATCCAGCTG 2565

QY 883 TTTTGGCATTCCTGCGACCGCACTGTTTCACTGGGCTTCAATTCCTCATCCAGCTG 942
Db 2566 TTTTGGCATTCCTGCGACCGCACTGTTTCACTGGGCTTCAATTCCTCATCCAGCTG 2625

QY 943 GCGGCGTTTTCGCGGATCCCTGCGGACCACTGTCGACTCGAGGCTTCTCTGATG 1002
Db 2626 GCTGGCGTTAGCCGCTTCCATGGGACCTGCTAACCACTCTGCGATCTTCTAGCG 2685
QY 1003 GCGGCGGCGTGGGCTTATGCTTGTGCGGACATCTGGCCAACTTACACTG 1059
Db 2686 GCGGCGATCGGCTGAGTATACGCACTGATCAACAGCCCACTTGCAGATTAATG 2742

Search completed: February 7, 2004, 19:25:45
Job time : 735.807 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:37:05 ; Search time 159.289 Seconds
(without alignments)
7088.094 Million cell updates/sec

Title: US-09-614-150A-8

Perfect score: 2558
Sequence: 1 99ccgctgccttccttc.....taatacaatacaagca 2558

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B COMB.seq:*
5: /cgn2_6/prodata/2/ina/5A COMB.seq:*
6: /cgn2_6/prodata/2/ina/5B COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	79	3.1	658	US-08-998-416-595	Sequence 595, App
2	67.8	2.7	2447	US-09-014-969-14	Sequence 14, Appl
3	67.6	2.6	240	US-08-528-417-6	Sequence 6, Appl
4	66.6	2.6	1696	US-09-835-811-1	Sequence 1, Appl
C 5	66	2.6	7218	US-08-232-463-14	Sequence 14, Appl
6	60.2	2.4	19124	US-08-487-826B-13	Sequence 13, Appl
7	59.8	2.3	3275	US-09-370-838-151	Sequence 151, App
8	58.8	2.3	6409	US-09-967-908A-1	Sequence 1, App
9	58.4	2.3	6124	US-08-213-419B-3	Sequence 3, Appl
10	58.2	2.3	1798	US-09-797-906-1	Sequence 1, Appl
11	57.6	2.3	1117	US-09-247-373B-33	Sequence 33, Appl
12	56.6	2.2	1447	US-09-443-041A-27	Sequence 27, Appl
13	55.4	2.2	1493	US-08-340-820-24	Sequence 24, Appl
14	55.4	2.2	1493	US-08-593-535-24	Sequence 24, Appl
15	54.8	2.1	2773	US-09-996-243-178	Sequence 178, App
16	53.8	2.1	3060	US-09-996-243-504	Sequence 504, App
17	53.6	2.1	2846	US-09-996-243-230	Sequence 230, App
C 18	53.4	2.1	602	US-08-764-100-8	Sequence 8, Appl
19	53.4	2.1	642	US-08-764-100-13	Sequence 13, Appl
C 20	53.4	2.1	643	US-08-764-100-7	Sequence 7, Appl
21	53.4	2.1	1048	US-09-489-847-38	Sequence 38, Appl
22	53.4	2.1	1813	US-09-071-224-3	Sequence 3, Appl
C 23	53.4	2.1	2993	US-08-764-100-2	Sequence 2, Appl
24	53.4	2.1	2993	US-08-764-100-10	Sequence 10, Appl
25	53.4	2.1	3000	US-08-764-100-9	Sequence 9, Appl
C 26	53.4	2.1	3001	US-08-764-100-1	Sequence 1, Appl
27	53.2	2.1	441	US-09-601-537-10	Sequence 10, Appl

C 28	53.2	2.1	615	US-08-998-416-186	Sequence 186, App
29	53.2	2.1	4121	US-09-601-537-9	Sequence 9, Appl
C 30	52.8	2.1	2565	US-09-252-991A-2277	Sequence 2277, Ap
31	52.8	2.1	3066	US-09-252-991A-2634	Sequence 2634, Ap
32	52.8	2.1	8920	US-08-446-855A-1	Sequence 1, Appl
33	52.8	2.1	8920	US-09-150-741-1	Sequence 1, Appl
C 34	52.6	2.1	1020	US-09-328-475C-43	Sequence 43, Appl
35	52.6	2.1	6243	US-09-056-075-1	Sequence 1, Appl
36	52	2.0	674	US-09-620-405B-465	Sequence 465, App
37	52	2.0	674	US-09-433-826B-465	Sequence 465, App
38	52	2.0	674	US-09-604-287A-465	Sequence 465, App
39	52	2.0	1474	US-08-821-994-64	Sequence 64, Appl
40	52	2.0	2634	US-09-463-238-3	Sequence 3, Appl
41	51.8	2.0	144	US-08-702-344-26	Sequence 26, Appl
42	51.8	2.0	2434	US-09-489-847-67	Sequence 67, Appl
43	51.2	2.0	1051	US-09-245-041-10	Sequence 10, Appl
44	51	2.0	2674	US-09-817-180-1	Sequence 1, Appl
45	51	2.0	6671	US-08-280-443-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-998-416-595/c
; Sequence 595, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jürgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38, 241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 595:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 658 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1408RP

US-08-998-416-595

Query Match 3.1%; Score 79; DB 3; Length 658;

Best Local Similarity 49.0%; Pred. No. 1.2e-09; Mismatches 220; Indels 0; Gaps 0;

Matches 211; Conservative 0; Mismatches 220; Indels 0; Gaps 0;

QY 2128 AAAAAACAAAAAATATGAAAAACACAGCAAGAACATTCATTCAGTCAATTA 2187

DB 647 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 588

QY 2188 AGCAATCCAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2247

DB 587 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 528

QY 2248 AGCATGTAATCCCTACACACACACACACACACACACACACACACACAC 2307

DB 527 AGCATATATATATATGTTTCAACAAATACATCTATATATATATATATAT 468

QY 2308 ATTTATTTATGATGAGGACGAGGATTTATTTATTTATTTATTTATTT 2367

DB 467 ACTATCTATGATATATATATATATATATATATATATATATATATATAT 408

QY 2368 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2427

DB 407 CATTAATAATAATAAGTATATATATATATATATATATATATATATAT 348

QY 2428 AGCAAG 2487

DB 347 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 288

QY 2488 CACGAAAGCCAAAGAAATTAATAATAATAATAATAATAATAATAATA 2547

DB 287 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 228

QY 2548 AATTAACAAGCA 2558

DB 227 AATAATAACAA 217

RESULT 2

US-09-014-969-14

Sequence 14, Application US/09014969

Patent No. 5965397

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

APPLICANT: McCoy, John M.

APPLICANT: Lavallee, Edward R.

APPLICANT: Racine, Lisa A.

APPLICANT: Metberg, David

APPLICANT: Treacy, Maurice

APPLICANT: Spaulding, Vikki

APPLICANT: Agostino, Michael J.

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/014,969

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.

REGISTRATION NUMBER: 41,323

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8284

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ. ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 2447 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-014-969-14

Query Match 2.7%; Score 67.8; DB 2; Length 2447;

Best Local Similarity 53.2%; Pred. No. 1.2e-06; Mismatches 123; Indels 0; Gaps 0;

Matches 141; Conservative 1; Mismatches 123; Indels 0; Gaps 0;

QY 2294 GTATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2353

DB 2160 GTATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2219

QY 2354 GCGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2413

DB 2220 ATGTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2279

QY 2414 AAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 2473

DB 2280 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 2339

QY 2474 AATTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2533

DB 2340 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 2399

QY 2534 AACGTAATTAACAATAATTAACAACA 2558

DB 2400 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 2424

RESULT 3

US-08-628-417-6

Sequence 6, Application US/08628417

Patent No. 5627054

GENERAL INFORMATION:

APPLICANT: GILLESPIE, DAVID

TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: U.S. ARMY CHEMICAL AND BIOLOGICAL

STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)

CITY: ABERDEEN PROVING GROUND

STATE: MARYLAND

COUNTRY: USA

ZIP: 21010-5423

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/628,417

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BIFFONI, ULYSSES J

REGISTRATION NUMBER: 39,908

REFERENCE/DOCKET NUMBER: DAM 398-94

TELECOMMUNICATION INFORMATION:

TELEPHONE: 410-671-1158

TELEFAX: 410-671-2534

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: P2730P1C13
;; CURRENT APPLICATION NUMBER: US/09/996,243
;;
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11

;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2004, 04:41:38 ; Search time 963.642 Seconds
(without alignments)
9778.248 Million cell updates/sec

Title: US-09-614-150A-8

Perfect score: 2558

Sequence: 1 ggcgcgcgcgccttccttc.....taatacaaatcaagca 2558

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	161.4	6.3	1559	9	US-09-908-419-1
2	161.4	6.3	1559	11	US-09-907-187-1
3	161.4	6.3	1559	13	US-10-056-790-1
4	161.4	6.3	1559	13	US-10-056-790-18
5	160	6.3	1376	9	US-09-908-419-5
6	160	6.3	1376	11	US-09-907-187-5
7	160	6.3	1376	13	US-10-056-790-5
8	121.6	4.8	1413	13	US-10-056-790-45
9	107.2	4.2	894	13	US-10-056-790-29
10	96.8	3.8	367378	13	US-10-312-841-2
11	95.2	3.7	1224	9	US-09-908-419-3
12	95.2	3.7	1224	11	US-09-907-187-3
13	95.2	3.7	1224	13	US-10-056-790-3
14	95.2	3.7	2118	12	US-10-108-260A-2042
15	94.8	3.7	6668	13	US-10-311-455-1670

C 16	93.6	3.7	1406	13	US-10-311-455-1931	Sequence 1931, Ap
C 17	87.8	3.4	14798	13	US-10-311-455-1005	Sequence 1005, Ap
C 18	86	3.4	9539	15	US-10-240-453-54	Sequence 54, Appl
C 19	86	3.4	9539	15	US-10-239-676-52	Sequence 52, Appl
C 20	85.6	3.3	529	10	US-09-983-965-2109	Sequence 2109, Ap
C 21	85.4	3.3	12237	13	US-10-311-455-2331	Sequence 2331, Ap
C 22	83.4	3.3	593	13	US-10-056-790-19	Sequence 19, Appl
C 23	83.4	3.3	8392	13	US-10-311-455-1463	Sequence 1463, Ap
C 24	82.6	3.2	6115	13	US-10-311-455-1774	Sequence 1774, Ap
C 25	82.4	3.2	367378	13	US-10-312-841-1	Sequence 1, Appli
C 26	82	3.2	6070	13	US-10-311-455-1652	Sequence 1652, Ap
C 27	82	3.2	6070	13	US-10-240-485-132	Sequence 132, Appl
C 28	82	3.2	9964	13	US-10-311-455-71	Sequence 71, Appl
C 29	81.6	3.2	17934	13	US-10-311-455-1692	Sequence 1692, Ap
C 30	81.2	3.2	6145	13	US-10-311-455-945	Sequence 945, Appl
C 31	81.2	3.2	6171	13	US-10-311-455-761	Sequence 761, Appl
C 32	80.8	3.2	6419	13	US-10-311-455-240	Sequence 240, Appl
C 33	80.6	3.2	7657	13	US-10-311-455-1995	Sequence 1995, Ap
C 34	80.6	3.2	7657	15	US-10-239-676-185	Sequence 185, Appl
C 35	80.2	3.1	12142	13	US-10-311-455-1646	Sequence 1646, Ap
C 36	80	3.1	7597	13	US-10-311-455-986	Sequence 986, Appl
C 37	79.6	3.1	446	10	US-09-960-352-1400	Sequence 3400, Ap
C 38	79.6	3.1	7118	13	US-10-311-455-803	Sequence 803, Appl
C 39	79.4	3.1	7571	13	US-10-311-455-500	Sequence 500, Appl
C 40	79.4	3.1	10369	13	US-10-311-455-365	Sequence 365, Appl
C 41	79	3.1	6175	13	US-10-311-455-1280	Sequence 1280, Ap
C 42	79	3.1	6668	13	US-10-311-455-1669	Sequence 1669, Ap
C 43	79	3.1	11996	13	US-10-240-485-45	Sequence 45, Appl
C 44	78.8	3.1	5070	13	US-10-311-455-774	Sequence 774, Appl
C 45	78.8	3.1	13574	13	US-10-311-455-1289	Sequence 1289, Ap

ALIGNMENTS

RESULT 1
US-09-908-419-1
; Sequence 1, Application US/09908419
; Patent No. US20020022029A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: HUMAN RRP SEQUENCES AND METHODS OF USE
; FILE REFERENCE: EX01-041C
; CURRENT APPLICATION NUMBER: US/09/908,419
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/219,289
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/277,487
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/277,471
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-419-1

Query Match 6.3%; Score 161.4; DB 9; Length 1559;
Best local Similarity 57.4%; Pred. No. 2.6e-28;
Matches 329; Conservative 0; Mismatches 241; Indels 3; Gaps 2;

QY	778	CTGTGCTATGCGCCGCGCGCTGCGAGTGTGCGCTTTTAACTATGCTTCTG	837
DB	730	CTTGTGTATGACACCCGCGGACCGTCCGCGCTGCTCTCACTAATGTTGATG	789
QY	838	CACGCCAATGTTTCACTGCTGCTTCAATATGCTATGCTGCTTTTGGCAATGCC	897
DB	790	CACGTTGGGCTGAGAGCTGAGGTTTCAAGCTCTCTGACCTGATGATCGGGTCC	849
QY	898	CTGAGAGTGATGACGCGGACGCGAGATGCGGCTGATCTAATGAGCGGCGTTTGGC	957

Db 850 CTGAGATGATGACGCGCCCTGCTCCGCATCAGCTCTTACCTGGCAGCGGTCTGGCA 909
Qy 958 GATTCCTGGGACACAGTGTCTGAGAGTCTCTCTGGTGGGCGCAGCGGTGGC 1017
Db 910 GGTCTCTTAACCTCTTCCATCACCGACATGCGGCGCCCGGTGTGGAGGCTCCGCGGG 969
Qy 1018 GTCTATGACCTGTGGCCGACATCTGGCCACATTACATGAATATATGCGACATGAG 1077
Db 970 GTCTAGCCCTGTGTCTGGACACCTGGCCAGTTGTATGAATCTGGGATGAGA 1029
Qy 1078 AGCGATCCACGCACTCGGATCCGTGTCTTGTCT-CTCGGATCTGGGCTATGC 1136
Db 1030 TGTCTTACAAAGTGTGAGAGATGTGCTGGCTGTGTGATGAGCTCCGAGTGGGC 1089
Qy 1137 TCTCTACACCAATCTTGTGATGAGAGCGCT--TGCCAAAGGTCCCAAGTGTGAC 1194
Db 1090 CGGGCGGTGTGGCTTCTTCTCCCGCGCTCCGCTCGGCGCCACAGCCAGCTTC 1149
Qy 1195 ATGCGCACCTGACGGGAGCCCTGGCAGACTAAGATCGGCTTCTGTGCTAAGAAC 1254
Db 1150 ATGGCGCACTGGCAGCGCGGTGTGGGATGAGATGGCTTGAACATCTTCCGAGC 1209
Qy 1255 TTGCGTACCGAGAGTACGAGAGCTTCTGTGGCTGAGCGTGGCGCTTACTGTGC 1314
Db 1210 TACGAGAGAGCGCTGGCGGACAGTGGGCTGGTGGGTGCTGCTGAGCTTACGCGACC 1269
Qy 1315 TTGACCGTCTTGGCCATCTTTCACCTGATC 1347
Db 1270 TTCTGCTCTTCCCGCTTCTTGGAACTTTC 1302

RESULT 2

US-09-907-187-1
; Sequence 1, Application US/0907187
; Publication No. US2003027144A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: HUMAN RRP SEQUENCES AND METHODS OF USE
; FILE REFERENCE: EX01-041C
; CURRENT APPLICATION NUMBER: US/09/907,187
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 60/219,289
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/277,487
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/277,471
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-187-1

Query Match 6.3%; Score 161.4; DB 11; Length 1559;
Best Local Similarity 57.4%; Pred. No. 2.6e-28;

Matches 329; Conservative 0; Mismatches 241; Indels 3; Gaps 2;

Qy 778 CTGCTATGCGCGGACCGGCGCTGACAGTGTGGCGCTTCTTATGCTACATGTTCTG 837
Db 730 CTGTGTACACACCCCGGACCGGTGCCCGCGCTTCTCTCACTACATGTTCTATG 789
Qy 838 CACGCAACTGTTTCCACTGGGCTTCAATATGTCATTCAGCTGTTTTTGGCATTCC 897
Db 790 CAGGTGTGGCTGAGAGAGCTGGGGTTCAACGCCCTCTGACAGTGAATGCGGGTGC 849
Qy 898 CTGAGATATGACAGGCAAGGCAAGATGCGGTATCTAATAGGCGGCGCTTTTGGC 957
Db 850 CTGGAATGTGACAGGCGCTGCTCCGATCAGCTGCTTACCTGGGAGGCGTGTGGCA 909
Qy 958 GATTCCTGGGACACAGTGTCTGACACTCGAGGTTCTCTGATGGGCGCAGCGGTGC 1017

Db 910 GGTCTCTTAACCTCTTCCATCACCGACATGCGGCGCCCGGTGTGGAGGCTCCGCGGG 969
Qy 1018 GTCTATGACCTGTGGCCGACATCTGGCCACATTACATGAATATATGCGACATGAG 1077
Db 970 GTCTAGCCCTGTGTCTGGACACCTGGCCAGTTGTATGAATCTGGGATGAGA 1029
Qy 1078 AGCGATCCACGCACTCGGATCCGTGTCTTGTCT-CTCGGATCTGGGCTATGC 1136
Db 1030 TGTCTTACAAAGTGTGAGAGATGTGCTGGCTGTGTGATGAGCTCCGAGTGGGC 1089
Qy 1137 TCTCTACACCAATCTTGTGATGAGAGCGCT--TGCCAAAGGTCCCAAGTGTGAC 1194
Db 1090 CGGGCGGTGTGGCTTCTTCTCCCGCGCTCCGCTCGGCGCCACAGCCAGCTTC 1149
Qy 1195 ATGCGCACCTGACGGGAGCCCTGGCAGACTAAGATCGGCTTCTGTGCTAAGAAC 1254
Db 1150 ATGGCGCACTGGCAGCGCGGTGTGGGATGAGATGGCTTGAACATCTTCCGAGC 1209
Qy 1255 TTGCGTACCGAGAGTACGAGAGCTTCTGTGGCTGAGCGTGGCGCTTACTGTGC 1314
Db 1210 TACGAGAGAGCGCTGGCGGACAGTGGGCTGGTGGGTGCTGCTGAGCTTACGCGACC 1269
Qy 1315 TTGACCGTCTTGGCCATCTTTCACCTGATC 1347
Db 1270 TTCTGCTCTTCCCGCTTCTTGGAACTTTC 1302

RESULT 3

US-10-056-790-1
; Sequence 1, Application US/10056790
; Publication No. US20030165497A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: RRP SEQUENCES AND KNOCKOUT MICE AND USES THEREOF
; FILE REFERENCE: RRPCTP2002
; CURRENT APPLICATION NUMBER: US/10/056,790
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US 09/908,419
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/219,289
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 60/277,487
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/277,471
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/304,863
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/305,017
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/328,491
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-056-790-1

Query Match 6.3%; Score 161.4; DB 11; Length 1559;
Best Local Similarity 57.4%; Pred. No. 2.6e-28;

Matches 329; Conservative 0; Mismatches 241; Indels 3; Gaps 2;

Qy 778 CTGCTATGCGCGGACCGGCGTGTGACAGTGTGGCGCTTCTTATGCTACATGTTCTG 837
Db 730 CTGTGTACACCCCGGACCGGTGCCCGCGCTTCTCTCACTACATGTTCTATG 789
Qy 838 CACGCAACTGTTTCCACTGGGCTTCAATATGTCATTCAGCTGTTTTTGGCATTCC 897

Db 595 CTGCTGCCCCCACCCTGTTCAATGATCAAGTCAGCTGAGGTTGCCCTTTTCTT 654
Qy 711 CTACAGACCGCTACACAAATGCCCCCGCAATTTCCGCTCAATCCGCTCGGA 770
Db 655 CTACAAATGGGGTGTCTAGTCAATTTGTACTGCAAGTACTCAATCACTTAA 714
Qy 771 TTCCGCTGCTCTATCCGCGGACCGCGCTGCAAGTGTGCGCTTTTAACTACAT 830
Db 715 GAACCTCCGTGTTTACACCCACAGCTGGAGACAGGTTTGGCGTACTGACATACAT 774
Qy 831 GTTCTGCAAGCCCACTGGTTTCACTGGGCTTCAATATGTCATCACTGTTTCTT 890
Db 775 CTTCATGATGCAAGGATAGAACCTTGGAATCAATGTGCTCTCAAGTGTGCTG 834
Qy 891 CATTCCTGAGAGTATGACAGGACCGGACAGATCCGCGTATCAATGAGCGGCGCT 950
Db 835 GGTGCCCCCTGGAATGTGATGAGTCAAGCCCAATTTGGCTTGTCTACGTTGCG 894
Qy 951 TTTTCCGATCCCTGAGGACCAAGTGTCTGCACTGCGAGTCTTCTGTTGGGCGCAG 1010
Db 895 TGTGCAAGGGTCTTGGCAGTGTCTGTGCTGACATGACGCTCCAGTGTGGGCTTTC 954
Qy 1011 CGGTGCGCTATGATCCCTGTTGGCGGACATCTGGCCAACTTACACTGAATATGCGCA 1070
Db 955 TGGAGGGGTGTATGCTCTGCTCTGCTGCTGCGCCATCTGCGCAATGTATGAATGTCAGG 1014
Qy 1071 CATGAAGAGCGATTCACAGCACTCGGATCCGTTGTCATCTTGTCTCCGCTGCGATGCGG 1130
Db 1015 CATGAAGTGCAGTTCAAGTGTCTGCGGATGAGTGTGCGCTTATCTGTATGACATGGA 1074
Qy 1131 CTATGCTCTCTACACC---CAATACTTGCATGAGAGCGCTTGGCCAAAGGTCGCCAGGT 1187
Db 1075 GTTTGGGGCGGGCGGTGTGCTCGGCTTCCACCGGTGGGCTTATCCCGGCGCTCAACC 1134
Qy 1188 GTCTGATATGCTCCCACTGACGGGAGCGCTGCGACAGATCAAGTGTGCTTCTGCTGT 1247
Db 1135 AAGCTTGTGGCCACACTGTGGGTGCGGTGCGCGGATCACCTGGGCGTGTGTGCTT 1194
Qy 1248 AAGAACTTGGTCAAGGAGATGACGAGCTCATCTGGTGGCTAGCGTTGGGCGCTCA 1307
Db 1195 GAGGAATCTACAGACAGAGGCTCCAGGACCACTGCTGTGATTTTGTGGCATGTA 1254
Qy 1308 CTGTGCTTCAACCGCTTTCGCCATGCTTTTCAACT 1343
Db 1255 CACCGTCTTGTGCTGTGCTGTCTTCTGGAACAT 1290

RESULT 6
US-09-907-187-5

; Sequence 5, Application US/0907187
; Publication No. US20030027144A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: HUMAN RBP SEQUENCES AND METHODS OF USE
; FILE REFERENCE: EX01-041C
; CURRENT APPLICATION NUMBER: US/09/907,187
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 60/219,289
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/277,487
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/277,471
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1376
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-187-5

Query Match 6.3%; Score 160; DB 11; Length 1376;
Best Local Similarity 52.9%; Pred. No. 5.1e-28;

Matches 368; Conservative 0; Mismatches 325; Indels 3; Gaps 1;
Qy 651 CCGCAGCACTGAGCCCTGTTATCTAGTATCTCAATGATGATTCATCTTCCG 710
Db 595 CTGCTGCCCCCACCCTGTTTATGATCAATCAAGTGTGAGGTTGCCCTTTTCT 654
Qy 711 CTACAGACCGCTACACAAATGCCCCCGCAATTTCCGCTCAATCCGCTCGGA 770
Db 655 CTACAAATGGGGTGTCTAGTCAATTTGTACTGCAAGTACTCAATCACTTAA 714
Qy 771 TTCCGCTGCTCTATCCGCGGACCGCGCTGCAAGTGTGCGCTTTTAACTACAT 830
Db 715 GAACCTCCGTGTTTACACCCACAGCTGGAGACAGGTTTGGCGTACTGACATACAT 774
Qy 831 GTTCTGCAAGCCCACTGGTTTCACTGGGCTTCAATATGTCATCACTGTTTCTT 890
Db 775 CTTCATGATGCAAGGATAGAACCTTGGAATCAATGTGCTCTCAAGTGTGCTG 834
Qy 891 CATTCCTGAGAGTATGACAGGACCGGACAGATCCGCGTATCAATGAGCGGCGCT 950
Db 835 GGTGCCCCCTGGAATGTGATGAGTCAAGCCCAATTTGGCTTGTCTACGTTGCG 894
Qy 951 TTTTCCGATCCCTGAGGACCAAGTGTCTGCACTGCGAGTCTTCTGTTGGGCGCAG 1010
Db 895 TGTGCAAGGGTCTTGGCAGTGTCTGTGCTGACATGACGCTCCAGTGTGGGCTTTC 954
Qy 1011 CGGTGCGCTATGATCCCTGTTGGCGGACATCTGGCCAACTTACACTGAATATGCGCA 1070
Db 955 TGGAGGGGTGTATGCTCTGCTCTGCTGCTGCGCCATCTGCGCAATGTATGAATGTCAGG 1014
Qy 1071 CATGAAGAGCGATTCACAGCACTCGGATCCGTTGTCATCTTGTCTCCGCTGCGATGCGG 1130
Db 1015 CATGAAGTGCAGTTCAAGTGTCTGCGGATGAGTGTGCGCTTATCTGTATGACATGGA 1074
Qy 1131 CTATGCTCTCTACACC---CAATACTTGCATGAGAGCGCTTGGCCAAAGGTCGCCAGGT 1187
Db 1075 GTTTGGGGCGGGCGGTGTGCTCGGCTTCCACCGGTGGGCTTATCCCGGCGCTCAACC 1134
Qy 1188 GTCTGATATGCTCCCACTGACGGGAGCGCTGCGACAGATCAAGTGTGCTTCTGCTGT 1247
Db 1135 AAGCTTGTGGCCACACTGTGGGTGCGGTGCGCGGATCACCTGGGCGTGTGTGCTT 1194
Qy 1248 AAGAACTTGGTCAAGGAGATGACGAGCTCATCTGGTGGCTAGCGTTTGGGCGCTCA 1307
Db 1195 GAGGAATCTACAGACAGAGGCTCCAGGACCACTGCTGTGATTTTGTGGCATGTA 1254
Qy 1308 CTGTGCTTCAACCGCTTTCGCCATGCTTTTCAACT 1343
Db 1255 CACCGTCTTGTGCTGTGCTGTCTTCTGGAACAT 1290

RESULT 7

US-10-056-790-5
; Sequence 5, Application US/10056790
; Publication No. US20030165497A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: RBP SEQUENCES AND KNOCKOUT MICE AND USES THEREOF
; FILE REFERENCE: RRP2002
; CURRENT APPLICATION NUMBER: US/10/056,790
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US 09/908,419
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/219,289
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 60/277,487
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/277,471
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/304,863
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05

;; PRIOR APPLICATION NUMBER: US 60/305,017
;; PRIOR FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: US 60/328,605
;; PRIOR FILING DATE: 2001-10-10
;; PRIOR APPLICATION NUMBER: US 60/328,491
;; PRIOR FILING DATE: 2001-10-10
;; NUMBER OF SEQ ID NOS: 70
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 5
;; LENGTH: 1376
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-056-790-5

Query Match 6.3%; Score 160; DB 13; Length 1376;
Best Local Similarity 52.9%; Pred. No. 5,1e-28;
Matches 368; Conservative 0; Mismatches 325; Indels 3; Gaps 1;

QY 651 CCGCCAGCACTGAGCCCTGCTTCACTTATGATCTTCATCATTTGAGATTGCCATCTTGGC 710
DB 595 CTGCTGCCCCCACCCTGCTTATGATCAAGTCAGCTGAGAGTTGCCCTTTTCT 654
QY 711 CTACAGCCGCTACACATATGCCCCCAGAAATTTGGGCTAACCCGTTCCGATTCGTCGA 770
DB 655 CTACAAATGGGCTGCTCACTGATCTTATGAGTACTGACAGTACTCATCCAGTTACTTGA 714
QY 771 TTCCGCTGCTGCTTATGAGCCGAGACCGGAGTCTGAGTGGCGCTTCTTAGTACAT 830
DB 715 GAACCTCCCTGCTTATACACCACAGCTGAGAGCAGTTTGGCGCTTACCTGACATACAT 774
QY 831 GTTCTGACAGCCAACTGCTTCCACCTGAGCTTCAATATGTCATCAGCTGTTTTTGG 890
DB 775 CTTCATGACATGAGGAGTGAACACCTGGAGCTCAATGAGTCTGAGCTGCTGAGTGG 834
QY 891 CATTCCTCTGAGAGTATGACAGGACCGGACAGATCGGCTGATTTAATGCGGGCGT 950
DB 835 GGTGCCCCCTGAGATGCTCATGAGACCCAGAAATGGGCTTGTCTACGTGGCCGCTGT 894
QY 951 TTTTCCCGGATCCCTGGGACACAGTGTGCTGAGCTTCTGAGAGTCTTCTGTTGGGCGCAG 1010
DB 895 TGTGAGAGGCTCTTGGCAGTGTCTGTGCTGACATGACCGCTCCAGTGTGGGCTTTC 954
QY 1011 CGGTGCTCTATGACCTGCTTGGCCGACATCTGACCAATTTACATTAAGTAATGCGCA 1070
DB 955 TGGAGGGGTATGATGCTGTCTGTCTGCTGCTGAGCAATTTGATTAATGACGTGCAAG 1014
QY 1071 CATGAAGACGATTCACGCAACTCGGATCCGTTGTCATCTTGTCTCTGCGATCTGG 1130
DB 1015 CATGAAGTCCAGTTTCAAGCTGCTGCGATGAGTGGCTGTGGCCCTTATCTGATGAGCATGGA 1074
QY 1131 CTATGCTCTTACACC---CAATACTTGCATGGAAGCGCTTGGCCAAAGGTCCCAAGT 1187
DB 1075 GTTTGGGCGGGCTGTGTGCTCCGCTTCCACCGTGGCTTATCCCCGTGGCCCTTACCC 1134
QY 1188 GTCTGATATGCTCCACTGACGAGGAGCCCTGGCAGAGACTAAGATCGGCTTTTGGTGT 1247
DB 1135 AAGCTTTGTTGGCGCACTTGGTGGCTGGTGGCCGTGGGCAATCACTTGGGCTGTGCT 1194
QY 1248 AAAAAGCTTCGATCAACGAGAGTACAGAGACTCATCTGTGTGCTAAGCTTGGGCGCTGA 1307
DB 1195 GAGGAACCTACGAGCAGAGGCTCAGAGACAGTACATGTGTGGATTTTGTGGCCATGTA 1254
QY 1308 CTGTGCTTACCGCTTTTGGCATGCTTTTCAACT 1343
DB 1255 CACCGCTTGTGCTGTGTGCTGTCTTTCGAAACAT 1290

RESULT 8
US-10-056-790-45
; Sequence 45, Application US/10056790
; Publication No. US20030165497A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.

;; TITLE OF INVENTION: RRP SEQUENCES AND KNOCKOUT MICE AND USES THEREOF
;; FILE REFERENCE: RRP0202
;; CURRENT APPLICATION NUMBER: US/10/056,790
;; CURRENT FILING DATE: 2002-01-23
;; PRIOR APPLICATION NUMBER: US 09/908,419
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: US 60/219,289
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: US 60/277,487
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: US 60/277,471
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: US 60/304,863
;; PRIOR FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: US 60/296,076
;; PRIOR FILING DATE: 2001-06-05
;; PRIOR APPLICATION NUMBER: US 60/305,017
;; PRIOR FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: US 60/328,605
;; PRIOR FILING DATE: 2001-10-10
;; PRIOR APPLICATION NUMBER: US 60/328,491
;; PRIOR FILING DATE: 2001-10-10
;; NUMBER OF SEQ ID NOS: 70
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 45
;; TYPE: DNA
;; LENGTH: 1413
;; ORGANISM: Mus musculus
US-10-056-790-45

Query Match 4.8%; Score 121.6; DB 13; Length 1413;
Best Local Similarity 54.6%; Pred. No. 1,1e-18;
Matches 309; Conservative 0; Mismatches 224; Indels 33; Gaps 2;

QY 778 CTGTCTATGCGCCGAGCCGCTGTCAGAGTGTGGCGCTTTTAACTACATGTTCTCTG 837
DB 835 CTGTGTATCACCCAGACACCGCTCTGGGATGAGCGCTTCTCACTCAATGTTTCAG 894
QY 838 CAGCCAACTGTTCCACCTGAGCTTCAATATGATCATCAGCTGTTTTTGGCAATGCC 897
DB 895 CATGTGGGCTGAGAGCTAGAGTTCAATGCTTCTGACATGATGATGCTGTGCTCC 954
QY 898 CTGAGAGTATGACAGGACCGGACAGATCGGCTGATCTACATGAGCGGCGCTTTTGGC 957
DB 955 CTGAGATGTATCATGCGTGTGTCGATCAAGCTGCTTACTGCGGCGGCTGTGCGCA 1014
QY 958 GATCTCTGGGACCAAGTGTCTGACATCTGAGAGTCTTCTGTGGGCGCCAGCGGTGTC 1017
DB 1015 GGTCTCTGACATGCTCTATCAAGATATGCGTGGCCCGGTGTAGGGGGCTCTGAGAGG 1074
QY 1018 GTCTATGCTTGTGGCCGACATCTGGCCAACTTACATGAATATGCGCATGAG 1077
DB 1075 GTCTATGCTTGTGCTCAGACACCTGGCAATGTT----- 1110
QY 1078 AAGCATCAACCACTCGATCCGTCATCTTGTCTCTGCGATCTGGGCTATGCT 1137
DB 1111 -----GTCTATGTATGGAATCTGCTTTCTGGGAGTGAATTCGAATGGGCGGCT 1164
QY 1138 CTCTACACCAATCTTGCATGAGAAAGCGCTTGGCCAAAGGTCCCAAGTGTCTACAT 1197
DB 1165 GTGT---GGCTACGCTTCTCCCACTACATGCTGCTCAGGCCCAAGCCAGCTTCAG 1221
QY 1198 GGCACCTGAGCGGAGCCCTGGCAGAGCTAAGATTCGCTTCTGTGTAAAGAACTTC 1257
DB 1222 GCACACTGGCTGTGCAAGTGTATGAGTATGAGCTTACATCTTCTGGAGATAT 1281
QY 1258 GGTACCGAGATACAGAGCTCATCTGTGGTGTAGCGTGTGGGCTCTACTGTGCTTC 1317
DB 1282 GAGGAACGCTGAGGACCAAGTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1341
QY 1318 ACCGTCTTGGCCATGCTTTTCAACT 1343
DB 1342 CTGCTTGTGCTCATCTTTCGAAACGT 1367

[illegible]

```

RESULT 11
US-09-908-419-3
; Sequence 3, Application US/09908419
; Patent No. US2002002209A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: 'HUMAN RRP SEQUENCES AND METHODS OF USE
; FILE REFERENCE: EX01-041C
; CURRENT APPLICATION NUMBER: US/09/908,419
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/219,289
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/277,487
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/277,471
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-419-3

```

Query Match	3.7%	Score 95.2	DB 9	Length 1224
Best Local Similarity	50.4%	Pred. No. 2.7e-12		
Matches	285	Conservative	0	Mismatches 278; Indels 3; Gaps 2
Qy	782	TCTATCGCCCGGACCGGCGCTTGACAGAGTGTGGCGCTTCTTTAGTACATAGTTTCCCTGCACG	841	
Dp	518	TCTACAGTCTTGAAGAAGAGGAGAGAGCGCTGAGGTTTATCTCATACATGCTGGTACATG	577	
Qy	842	CCAACTGAGTTCACCTGGGCTTCAATTATGTCATCCAGCTGTGTTTTTGGCATTTCCCTGG	901	
Dp	578	CTGAGAGTTGACGACATCTTTGGGGAGATCTTTGTATGACAGCTGTGTTTTGGGTAATCCCTGG	637	
Qy	902	AGGTGATGACAGGACAGCGCAGGAATCGGCGTATCTTAACATGGCGGGCGTTTTTGGCGGAT	961	
Dp	638	AAATGTGTCACAAAGAGCCCTCGGTGTGGGGCTGTGTACTGTGCAGAGATGATTCAGAGGGT	697	
Qy	962	CCCTGGGACACAGTGTGTCGACATCGGAGGTCTTCTGTGGGGCGCCAGCGGTGGCGTCT	1021	
Dp	698	CCCTTGCCAGCTCCATCTTTGACCCACTCATGATATCTGTGGGAAGCTTTCAGAGGAGTCT	757	
Qy	1022	ATGCCCTTGTGGCCGACATCGGCGCAACATTACATTAACATATGCGACATGGAAGACG	1081	
Dp	758	ATGCTCTATGGAGGCTATTTTATGAATGTTCTGTGTATTTTCAAGAAATGATTTCTCG	817	
Qy	1082	CATCCAGCAACTCGGATTCGTTGATCTTGTGTCTCTCGAGATCTGGAGTATGCTCTCT	1141	

Db	818	CTTTGGAAATTTTCAGACTGCTGATCATTCAATCCGATATATGTGTGGAC-ATGGGANTTT	876
Qy	1142	ACACCCAAATCTTGATGAGAAAGCGCCTTTCGCCAAGGGTCCCAAGGTGTGTAACATTGGCCC	1201
Db	877	GCCTCTAATGAAGAGTTCTTTGTTC - -GAAGATGAGTCTCCGGTGTCTTTTGCAAGTC	934
Qy	1202	ACCTGACGGGAGCCCTGGCAGAGCTAACGATTCGGCTTTCTGTCGCTAAGAACACTTCGGTC	1261
Db	935	ACATTGAGGCGATTTGCTGGAATGTCCATTGGCTCACAGGTTTAAAGTCGCTTATTA	994
Qy	1262	ACGAGAGTACGAGCAGCTATCTGAGGGCTAGGAGGTGGGGGTCTACTGTGTCCTTCAACCG	1321
Db	995	AAGACATGCGAAGAATCCAAAGTTTGGATAGCAATTCGCAATATTAGCTGTGTGCT	1054
Qy	1322	TCCTTGGCATTGTTTCAACCTATC	1347
Db	1055	TATTTGCTGTGTTTTCACACTTTTC	1080

```

RESULT 12
US-09-907-187-3
? Sequence 3, Application US/09907187
? Publication No. US20030027144A1
? GENERAL INFORMATION:
? APPLICANT: EXELIXIS, INC.
? TITLE OF INVENTION: HUMAN RRP SEQUENCES AND METHODS OF USE
? FILE REFERENCE: EX01-041C
? CURRENT APPLICATION NUMBER: US/09/907,187
? CURRENT FILING DATE: 2001-07-17
? PRIOR APPLICATION NUMBER: 60/219,289
? PRIOR FILING DATE: 2000-07-19
? PRIOR APPLICATION NUMBER: 60/277,487
? PRIOR FILING DATE: 2001-03-21
? PRIOR APPLICATION NUMBER: 60/277,471
? PRIOR FILING DATE: 2001-03-21
? NUMBER OF SEQ ID NOS: 7
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 3
? LENGTH: 1224
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-907-187-3

```

Query Match	3.7%	Score 95.2	DB 11	Length 1224		
Best Local Similarity	50.4%	Pred. No. 2.7e-12				
Matches 285	Conservative 0	Mismatches 278	Indels 3	Gaps 2		
Qy	782	TCATCGCGCGAACC	CGCTCTGAGGTG	GGCGCTCTTAACTAATGTTCC	TGACG 841	
Db	518	TCATCAAGTCTCGA	MAAGAGGAGAAC	CTGAGAGTTATCTCA	TACATGCTGTGTAATG 577	
Qy	842	CCAACTGCTTCC	ACCTGGGCTTCA	TATACGATCAG	CGCTTTTGGCATTC	CCCGG 901
Db	578	CTGAGATTGACGA	CATCTTGGGAGAT	TTTGTATGACGCT	TTTGTGGATATCC	CTTGG 637
Qy	902	AGGTGATGCA	CGGCAAGG	CCAGATCG	ACGATCTACATG	CGCGGCTTTTGGCCGAT 961
Db	638	AAATGTCCAA	AAAGGCTCTCGTGTGG	GGCGCTGTGTA	CTGGCAGGAGTATG	CAAGGT 697
Qy	962	CCCTGGGACAC	AGTGTGTCG	ACATCGGAG	CTTCTCGTGTGG	CGCAGCGGTGGCTCT 1021
Db	698	CCCTTGGCAG	CTCCATTTTGA	CCCACTCGA	TATCTTGTGGAG	CTTCAGAGAGATCT 757
Qy	1022	ATGCGCTGTGG	CGGCACATCTGG	CCAACTTAC	TACTGAATTATG	CGCACATGAAGCG 1081
Db	758	ATGCTCTGA	TGGAGGCTATTTAT	TGAATGTTCTGG	GAATTTTCA	AGAAATGATTCCTG 817
Qy	1082	CATTCACGCA	ACTGGATTCG	TGTCACTT	TGTCTCTCGAT	CTGGGCTAATCTCTCT 1141
Db	818	CTTTTGAAT	TTTTTCAG	CTGCTAT	CACTCACTGA	TAAATTTGTGTGGAC-ATGCGATTT 876
Qy	1142	ACACCCAA	TACTTGATG	AAGCGCTT	TGGCAAGG	TCCCCCAGGTGCTCTACATTCGCC 1201

DB 877 GCTCTATATGAAAGTTCCTTTGTTCCCT--GAGATGGGTCCTCGGTCGTCCTTTTGACGCTC 934
QY 1202 ACCGACGGGAGCCCTGGGAGGACTAACGATCGGCTTTGCGTCTTAAAGACCTGGGTC 1261
DB 935 ACATTGAGGTGATGATTTGCTGGAAATGTCATGCTACACGAGTGTGAGCTTGATGATA 994
QY 1262 ACCGAGGTACGACGACCTCATCTGCTAGCTGAGGCGTCTACTGTGCTTCACCG 1321
DB 995 AAGCATGCTGAAAGATCCAGGTTTGGATTAGCAATTTGTCATATTAGCTTGTCT 1054
QY 1322 TCTTCGCCATCGTTTTCACCTGATC 1347
DB 1055 TATTGCTGTGTTTTCACATTTTC 1080

RESULT 13
US-10-056-790-3
; Sequence 3, Application US/10056790
; Publication No. US20030165497A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: RRP SEQUENCES AND KNOCKOUT MICE AND USES THEREOF
; FILE REFERENCE: RRP2002
; CURRENT APPLICATION NUMBER: US/10/056,790
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US 09/908,419
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/219,289
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 60/277,487
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/277,471
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/304,863
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/236,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/305,017
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/328,491
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-056-790-3

Query Match 3.7%; Score 95.2; DB 13; Length 1224;
Best Local Similarity 50.4%; Pred. No. 2,7e-12;
Matches 285; Conservative 0; Mismatches 278; Indels 3; Gaps 2;

DB 782 TCTATCGCCGACCGGCGCTCTGAGGTGCGGCTTCTTAAGTAAATGTTCTCTGACG 841
DB 518 TCTACAGTCTCTGAAAGAGGAGGAGGAGGCTTATCTCATATCATCTGTAATG 577
QY 842 CCAACTGTTTCCACCTGGGCTTCAATATCGTATCATGCTGTTTGGATTCCTCCG 901
DB 578 CTGAGATTACGACATCTTGGGGAATCTTTGTATGACGCTTGTGTTGGATTCCTCC 637
QY 902 AGGTGATGACGCGACGCGAGATCGGCTGATCTAATCATGCGCGGCTTTTGGCGGAT 961
DB 638 AATATGTCACAAGGCTCCGTTGAGGCTGAGTACCTGGCAGAGTATGACAGGT 697
QY 962 CCTTGGACACAGTGTGTGATCTGAGAGGCTTCTCTGTTGGGCGGCAAGGCTGCTT 1021
DB 698 CCTTGGACAGCTCATCTTTGACCCACTCAGATATCTTGTGGAGCTTCAGAGAGATCT 757
QY 1022 ATGCCCTGTTGGCGCACATCTGGCCAAATTAAGTAACTAATGCGCATGAGAGCG 1081

DB 758 ATGCTCTATGAGGAGCTTATTTATGAAATGTTCTGTGTAATTTTCAAGAAATGATCTCG 817
QY 1082 CATTCAGGAACTCGGATCCGTTGTGATCTTTGTCTCTCGCATGCGGCTAATGCTCTCT 1141
DB 818 CCTTGGAAATTTTCAAGCTGCTGATCATCTGATTAATGTGTGAC-ATGGGATTT 876
QY 1142 ACACCAATATCTCGATGAGAGCGCTTGGCCAAAGGTCCTCCAGTGTGTCATTTGCC 1201
DB 877 GCTCTATATGAAAGTTCCTTTGTTCCCT--GAGATGGGTCCTCGGTCGTCCTTTTGACGCTC 934
QY 1202 ACCGACGGGAGCCCTGGGAGGACTAACGATCGGCTTTGCGTCTTAAAGACCTGGGTC 1261
DB 935 ACATTGAGGTGATGATTTGCTGGAAATGTCATGCTACACGAGTGTGAGCTTGATGATA 994
QY 1262 ACCGAGGTACGACGACCTCATCTGCTAGCTGAGGCGTCTACTGTGCTTCACCG 1321
DB 995 AAGCATGCTGAAAGATCCAGGTTTGGATTAGCAATTTGTCATATTAGCTTGTCT 1054
QY 1322 TCTTCGCCATCGTTTTCACCTGATC 1347
DB 1055 TATTGCTGTGTTTTCACATTTTC 1080

RESULT 14
US-10-108-260A-2042
; Sequence 2042, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2042
; LENGTH: 2118
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2042

Query Match 3.7%; Score 95.2; DB 12; Length 2118;
Best Local Similarity 50.4%; Pred. No. 4e-12;
Matches 285; Conservative 0; Mismatches 278; Indels 3; Gaps 2;

QY 782 TCTATCGCCGACCGGCGCTCTGAGGTGCGGCTTCTTAAGTAAATGTTCTCTGACG 841
DB 1341 TCTACAGTCTCTGAAAGAGGAGGAGGAGGCTTATCTCATATCATCTGTAATG 1400
QY 842 CCAACTGTTTCCACCTGGGCTTCAATATCGTATCATGCTGTTTGGATTCCTCCG 901
DB 1401 CTGGAATTACGACATCTTGGGGAATCTTTGTATGACGCTTGTGTTGGATTCCTCC 1460
QY 902 AGGTGATGACGCGACGCGAGATCGGCTGATCTAATCATGCGCGGCTTTTGGCGGAT 961
DB 1461 AATATGTCACAAGGCTCCGTTGAGGCTGAGTACCTGGCAGAGTATGACAGGT 1520
QY 962 CCTTGGACACAGTGTGTGATCTGAGAGGCTTCTCTGTTGGGCGGCAAGGCTGCTT 1021
DB 1521 CCTTGGACAGCTCATCTTTGACCCACTCAGATATCTTGTGGAGCTTCAGAGAGTCT 1580
QY 1022 ATGCCCTGTTGGCGCACATCTGGCCAAATTAAGTAACTAATGCGCATGAGAGCG 1081
DB 1581 ATGCTCTGATGAGGAGCTATTTATGAAATGTTCTGTGAAATTTTCAAGAAATGATTCCTG 1640
QY 1082 CATTCAGGAACTCGGATCCGTTGTGATCTTTGTCTCTCGCATGCGGCTAATGCTCTCT 1141
DB 1641 CCTTGGAAATTTTCAAGCTGCTGATCATCTGATTAATGTGTGAC-ATGGGATTT 1699
QY 1142 ACACCAATATCTCGATGAGAGGCTTGGCCAAAGGTCCTCCAGTGTGTCATATTGCC 1201
DB 1700 GCTCTCTATGAAAGTTCCTTTGTTCCCT--GAGATGGGTCCTCGGTCGTCCTTTTGACGCTC 1757

QY 1202 ACCGAGGAGCCCTGGCAGAGCTAACGATCGCTTTCTGCTGAAGAACTTGGCTC 1261
DB 1758 ACATTCAGGTGATTTGGCTGAATGCTCATTTGGCTAACAGGCTTTAGCTGTTGATA 1817
QY 1262 ACCGAGAGTACGACGAGCTCATTTGGCTGCTAGCGCTTGGCGCTTACTGTGCTTACCG 1321
DB 1818 AAGCAGCTGCTGAAGAAAGTCCAAAGCTTTTGAATGCAATGCTGCTATTTAGCTTGTCT 1877
QY 1322 TCTTGCCATCGCTTTCAACCTGATC 1347
DB 1878 TATTGCTGTGTTTTCACATTTTC 1903
RESULT 15
US-10-311-455-1670/C
; Sequence 1670, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1670
; LENGTH: 6668
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 1936
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1670
Query Match 3.7%; Score 94.8; DB 13; Length 6668;
Best Local Similarity 45.4%; Pred. No. 1.1e-11;
Matches 425; Conservative 0; Mismatches 502; Indels 9; Gaps 2;
QY 1623 TTTCAGTCCGAACATAGTACCAAACTCAAAAAAAAAAAAAAAAACAAAATCAGAGAAAT 1682
DB 3655 TTTCATATATAAATTAACGTAAATATAAAGCAAAAAAAAAAAATTAATTTCTA 3596
QY 1683 CACTGAGCAAAAAGAAAGTSCGAGAGTGAAGAACATTAACCGAAGCCGAAACGTGTA 1742
DB 3595 CACTAAAAAAATTAATAAATCCCAAAATCCCGTACAAATACCGCTACAAATTTAAAA 3536
QY 1743 AACAAATGTTGTATGAACCAAGAAGCTGAATTTATTCGCTGTAAAAAACAAGTAAAA 1802
DB 3535 CTCATAACGTAAATTTAATTTAAAA--TAAATACATCAAAATTCAAAAAAGAACTTC 3478
QY 1803 ATCAAGAGAAATCAAGAGAGAAACAGAACTAATCGCTTCGCTATGATTTAAATG 1862
DB 3477 GTTAAACCTAAATCAAAATTAATTAATCAAAACATATCTTTGAAAAATTAATAAAATC 3418
QY 1863 ATTATCAATGTTTCAATTAATGTTTCTGTTTCTTAATTAATTAATTAATGATTCGCC 1922
DB 3417 AAATATTCATATTTAAAAATTCACGACTTCTAATCCCTTAATTAATTAATTAACAAAA 3358
QY 1923 GCAATTAAGTAAAGTAAATGAAATGCAAGAGTACAGAACTGTATCAAAATGTTTATACA 1982
DB 3357 AA 3298

QY 1983 TCCATTAAGCATATGCTCCGAAATTTAGATTAGTGTATTAATTAATATATTAG 2042
DB 3297 AA 3238
QY 2043 GTATTAATGACCTCTCTAACTAATTTGTTCAATTTGTAAATCTAATTAAGTGCACACTA 2102
DB 3237 AA 3178
QY 2103 GTCAACAAACAAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 2162
DB 3177 AA 3118
QY 2163 AGAACATTCATTTCAATCAATCAATTAAGCAATTCAGTAAATTAATTAATTAATTAATTA 2222
DB 3117 AA 3058
QY 2223 AGTCACTTAATGCTTACAAATTCAGCAATTCGTAATCCCTACACACACACACACACACA 2282
DB 3057 AA 2998
QY 2283 CACACTCGAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2342
DB 2997 AA 2938
QY 2343 TCGTCAATGAGCAACTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2402
DB 2937 AA 2885
QY 2403 AAACAGACGAGAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2462
DB 2884 AA 2825
QY 2463 TAAAAATATCCATTTGAAAAATTAACACAGAAAGCCAAAGAAATTAATAATCAAAACT 2522
DB 2824 AA 2765
QY 2523 TTCAGAAATACAGTAAATTAACAAATTAACAGCA 2558
DB 2764 AA 2729

Search completed: February 8, 2004, 20:44:40
Job time : 974.642 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:36:25 / Search time 5999.41 Seconds
(without alignments)
10362.824 Million cell updates/sec

Title: US-09-614-150A-8
Perfect score: 2558
Sequence: 1 ggcgcgcgccttccttc.....taatacaaatcaagcaa 2558

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pin:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_nam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_pbg:*
27: em_gse_vrl:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	643	25.1	643	10	AW941611 LD06131.3
2	618	24.2	634	12	B1372537 RE59529.5
3	514	20.1	514	9	AA247028 LD06131.5
4	325.4	12.7	525	12	B1615764 RH45513.5

5	325.4	12.7	528	12	B1592257	B1592257 RH09903.5
6	325.4	12.7	644	12	B1592434	B1592434 RH10122.5
7	325.4	12.7	644	12	B1629788	B1629788 RH58702.5
c 8	155	6.1	1191	29	CNS00707	AL061341 Drosophila
c 9	154	6.0	1101	29	CNS0027C	AL061341 Drosophila
c 10	152.4	6.0	1101	29	CNS00E18	AL069225 Drosophila
11	147.6	5.8	580	10	BF491389	BF491389 AT28124.5
12	135.4	5.3	820	14	CA316643	CA316643 UI-M-FW0-
13	135.4	5.3	834	14	CA31790	CA31790 UI-M-FW0-
14	130.8	5.1	799	12	B1544995	B1544995 603242466
15	123.8	4.8	712	13	B0454587	B0454587 603772383
16	120.8	4.7	838	12	B1822562	B1822562 603035272
17	119.4	4.7	737	13	B0454451	B0454451 603771504
18	117.6	4.6	627	10	BF781398	BF781398 602104581
c 19	113.8	4.4	1426	29	CC231597	CC231597 CH261-36A
c 20	112.8	4.4	598	12	BM490251	BM490251 PSP2n-PK0
c 21	111.6	4.4	668	12	B1289526	B1289526 UI-R-DK0-
c 22	111.6	4.4	670	12	B1289575	B1289575 UI-R-DK0-
c 23	111.4	4.4	460	9	AW325810	AW325810 16869 MAR
c 24	111	4.3	1074	29	B2696936	B2696936 SP-Ba009
c 25	110	4.3	1286	29	CC240677	CC240677 CH261-82I
c 26	109.8	4.3	710	14	CA319248	CA319248 UI-M-FW0-
c 27	109.8	4.3	1716	29	CC222065	CC222065 CH261-11A
c 28	109	4.3	1036	29	CNS03LWU	AL250012 Tetraodon
c 29	109	4.3	1352	29	CC268594	CC268594 CH261-67F
c 30	107.8	4.2	687	10	BR420745	BR420745 HMM002-B0
c 31	107.6	4.2	1885	9	AL514935	AL514935 AL514935
c 32	107.6	4.2	1042	29	CNS0148K	AL103838 Drosophila
c 33	107.2	4.2	600	10	BE976898	BE976898 b657d06.Y
c 34	107.2	4.2	894	10	BE778475	BE778475 601463932
c 35	107.2	4.2	1434	29	CC187638	CC187638 CH261-98P
c 36	106.8	4.2	794	29	CNS00962	AL053102 Drosophila
c 37	106.6	4.2	1226	14	CD504687	CD504687 CD469-B02
c 38	105.8	4.1	1596	29	CC293489	CC293489 CH261-62N
c 39	105	4.1	954	28	BH162327	BH162327 ENR1J37TR
c 40	104.2	4.1	777	29	CNS025WB	AL182612 Tetraodon
c 41	104	4.1	1101	29	CNS00K65	AL077453 Drosophila
c 42	102.2	4.0	899	13	BK453223	BK453223 BX453223
c 43	101.8	4.0	720	14	CB325245	CB325245 UI-R-D20-
c 44	101.2	4.0	1637	29	CC293588	CC293588 CH261-62P
c 45	101	3.9	888	29	CNS027EX	AL184578 Tetraodon

ALIGNMENTS

RESULT 1
AW941611/c 643 bp mRNA linear EST 23-APR-2001
LOCUS LD06131.3prIME LD Drosophila melanogaster embryo Bluescript
DEFINITION Drosophila melanogaster cDNA clone LD06131 3 similar to X52454:
D.melanogaster rho gene, mRNA sequence.
ACCESSION AW941611 GI:8117057
VERSION AW941611.1
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 643)
Harvey, D., Brokslein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G.M.
BDGP/HMMI Drosophila EST Project
Unpublished
Other ESTs: LD06131.5prIME
COMMENT Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or

|||||
241 CTAAGACATTGGATTACAAAGAACCCAGCATTTTGGATTATTAACATTGGACAGGC 300
|||||
327 AGAAAAACCTAAGATTTCTTCAACGGCCGACATGAGAACTTAACGAGATGTAA 386
|||||
301 AGAAAAACCTAAGATTTCTTCAACGGCCGACATGAGAACTTAACGAGATGTAA 360
|||||
387 CGAACAAGTGGATTGGGCCAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 446
|||||
361 CGAACAAGTGGATTGGGCCAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 420
|||||
447 TGGCAACCGCCGCAAGAGACCATCATTTGACATTTCCCGCGCTGCTCCAGTTCTTCAA 506
|||||
421 TGGCAACCGCCGCAAGAGACCATCATTTGACATTTCCCGCGCTGCTCCAGTTCTTCAA 480
|||||
507 CTCCTGCTGTACAGACCGCATTTGACAGACGCGAGAGAGACCTGCTGCAACCGCCCAAG 566
|||||
481 CTCCTGCTGTACAGACCGCATTTGACAGACGCGAGAGAGACCTGCTGCAACCGCCCAAG 540
|||||
567 CGAGCAATCTACATGACCAACGCGAGGCAATCCCGGCAAGCCACTTCCGAGTCCGAGGA 626
|||||
541 CGAGCAATCTACATGACCAACGCGAGGCAATCCCGGCAAGCCACTTCCGAGTCCGAGGA 600
|||||
627 TATCGCCTGCTGAGTACGTCACCGCCAGCAC 660
|||||
601 TATCGCCTGCTGAGTACGTCACCGCCAGCAC 634
|||||
RESULT 3 514 bp mRNA linear EST 23-APR-2001
AA247028
LOCUS LD06131.5prIME LD Drosophila melanogaster embryo Bluescript
DEFINITION Drosophila melanogaster cDNA clone LD06131 5 similar to rho:
FR0001004 'signal transduction' located on: 3L 62A3-62A3.1;
04/10/2001, mRNA sequence.
ACCESSION AA247028 GI:13766053
VERSION AA247028
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 514)
REFERENCE Harvey D., Brokstein P., Hong L., Evans-Holm M., Su C., Tsang G.,
Lewin S. and Rubin G.M.
BDGP/HMI Drosophila EST Project
TITLE Unpublished
JOURNAL On Jan 24, 1997 this sequence version replaced gi:1878347.
COMMENT Other ESTs: LD06131.3prIME
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AB003471: arm:3L [1212758,1513165]
estimated-cyto:61F4-62A6: 04/10/2001
Plate: LD.61 row: C column: 7
High quality sequence stop: 509.
Location/Qualifiers
1..514
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="LD06131"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="SOLR"
/clone_lib="LD Drosophila melanogaster embryo Bluescript"
/note="Organ: embryo; Vector: Bluescript SK; Site 1: EcoRI
; Site 2: XhoI; Constructed using Stratagene ZAP-cDNA
Synthesis kit. Oligo dt-primed and directionally cloned at

BASE COUNT 174 a 117 c 128 g 95 t
ORIGIN
Query Match 20.1%; Score 514; DB 9; Length 514;
Best Local Similarity 100.0%; Pred. No. 1.2e-57;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
|||||
44 CGAGCCGAGATCATCAAACTGAGAAAGTCGACTCGACTGAAACTGAAATTTGAACT 103
|||||
1 CGAGCCGAGATCATCAAACTGAGAAAGTCGACTCGACTGAAACTGAAATTTGAACT 60
|||||
104 GAAAGAAAGAAATTTCAATTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 163
|||||
61 GAAAGAAAGAAATTTCAATTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
|||||
164 AGAATTTCTGATACAGCTCTGTGATTTACGAGCGCAAACTAAGTTACCAATGTGG 223
|||||
121 AGAATTTCTGATACAGCTCTGTGATTTACGAGCGCAAACTAAGTTACCAATGTGG 180
|||||
224 AGCCGAAAAAGGAGTGAAGAAAGTCGCAATATGCGAGTAACTAAAGCATTTGATTT 283
|||||
181 AGCCGAAAAAGGAGTGAAGAAAGTCGCAATATGCGAGTAACTAAGCATTTGATTT 240
|||||
284 ACAAGAAACCCAGCATTTTGGATTATTAACATTGGACAGCGAGAAAACTTAAGATT 343
|||||
241 ACAAGAAACCCAGCATTTTGGATTATTAACATTGGACAGCGAGAAAACTTAAGATT 300
|||||
344 TCTTCAACGCGCGCAGCATGAGAACTTAACGCAATGTAAACGAAACCAAGTGGATT 403
|||||
301 TCTTCAACGCGCGCAGCATGAGAACTTAACGCAATGTAAACGAAACCAAGTGGATT 360
|||||
404 TGGGCGAGAGAAAGAGAGAGGCGTGCAGAGAGAGAGAGATGCAACCGCCGCAAG 463
|||||
361 TGGGCGAGAGAGAGAGAGGCGTGCAGAGAGAGAGAGATGCAACCGCCGCAAG 420
|||||
464 AGACATCATTTGACATTTCCGCGCGCTGCTCAATCTCTCAATCTCTGCTGTACAGA 523
|||||
421 AGACATCATTTGACATTTCCGCGCGCTGCTCAATCTCTCAATCTCTGCTGTACAGA 480
|||||
524 CGATTTGACAGACGCGAGAGACCTGCTGCAC 557
|||||
481 CGATTTGACAGACGCGAGAGACCTGCTGCAC 514
|||||
RESULT 4 525 bp mRNA linear EST 07-SEP-2001
BI615764
LOCUS RH45513.5prIME RH Drosophila melanogaster normalized Head pFLC-1
DEFINITION Drosophila melanogaster cDNA clone RH45513 5 similar to rho:
FR0001004 GO: [integral plasma membrane protein (GO:0005887);
Plasma membrane (GO:0005886); signal transduction (GO:0004871)]
located on: 3L 62A3-62A3.1; 08/18/2001, mRNA sequence.
ACCESSION BI615764 GI:15511289
VERSION BI615764
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 525)
REFERENCE Stapleton M., Brokstein P., Hong L., Tyler D., Berman B., Carlson
J., Champe M., Chavez C., Dorsett V., Farfan D., Fries E., George
R., Gonzalez M., Guarin H., Harris N., Li P., Liao G., Miers S.,
Munhall C. J., Nunoo J., Pacleb J., Paragas V., Park S.,
Phonanaavong S., Wan K., Yu C., Lewis S. B., Ceiniker S. and Rubin
G.M.
BDGP/HMI RH Drosophila EST Project
TITLE Unpublished
JOURNAL On Jan 24, 1997 this sequence version replaced gi:1878347.
COMMENT Other ESTs: LD06131.3prIME
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab

KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 644)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mista, S.,
Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
Phonananavong, S., Wan, K., Yu, C., Lewis, S.E., Celnikier, S. and Rubin
G.M.
BDGP/HMI RH Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AB003471: arm:3L [1212758,1513169]
estimated-cyto:61F4-62A6: 08/16/2001
Plate: RH.101 row: B column: 10
High quality sequence stop: 597.
Location/Qualifiers
1..644
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RH10122"
/sex="male and female"
/dev_stage="Adult"
/lab_host="DHS-alpha Tona"
/clone_lib="RH Drosophila melanogaster normalized Head
pLC-1"
/note="Organ: head; Vector: pFL1; Site 1: XhoI; Site 2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 232 a 136 c 144 g 132 t
ORIGIN
Query Match 12.7%; Score 325.4; DB 12; Length 644;
Best Local Similarity 99.4%; Pred. No. 3.9e-33;
Matches 337; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GGCCGGCTCCCTTTCGAAAGTCAGTTGGCGGCGCGAGCGGAGATCATCA 60
DB 103 GGCCGGCTCCCTTTCGAAAGTCAGTTGGCGGCGCGAGCGGAGATCATCA 162
QY 61 AACTGGAAGTCGACTCGACTCGAACTGAATTGAACCTGAAAGAAAGAAATATTC 120
DB 163 AACTGGAAGTCGACTCGACTCGAACTGAATTGAACCTGAAAGAAAGAAATATTC 222
QY 121 AAATTGTCGTGTGTGGTGAAGCAAGATATATCTCAAGATATCTGAATACAA 180
DB 223 AAATTGT-GTGTGTGTGGTGAAGCAAGATATATCTCAAGATATCTGAATACAA 281
QY 181 GCTCTGATTTACGAGCAGCAAACTAAGTTACCAATGTGCGAGCGGAAAAAGGAGT 240
DB 282 GCTCTGATTTACGAGCAGCAAACTAAGTTACCAATGTGCGAGCGGAAAAAGGAGT 341
QY 241 GAAAAACGTGCGAATATGCGAGCTAATCTAAAGCATTTGGATTACAAGAAACCGACGAT 300
DB 342 GAAAAACGTGCGAATATGCGAGCTAATCTAAAGCATTTGGATTACAAGAAACCGACGAT 401
QY 301 TTGGATTAAACATTTGCGAGCAGGAAAAACCTAAG 339
DB 402 TTGGATTAAACATTTGCGAGCAGGAAAAACCTAAG 440

RESULT 7

B1629788
LOCUS B1629788 644 bp mRNA linear EST 10-SEP-2001
DEFINITION RH58702.5prime RH Drosophila melanogaster normalized Head pFLC-1
Drosophila melanogaster cDNA clone RH58702.5 similar to rho:
Fban001004 GO: [integral plasma membrane protein (GO:0005877);
plasma membrane (GO:0005866); signal transduction (GO:0004871)]
located on: 3L 62A3-62A3; 08/23/2001, mRNA sequence.
B1629788
B1629788.1 GI:15531998
EST.
Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 644)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mista, S.,
Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
Phonananavong, S., Wan, K., Yu, C., Lewis, S.E., Celnikier, S. and Rubin
G.M.
BDGP/HMI RH Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AB003471: arm:3L [1212758,1513169]
estimated-cyto:61F4-62A6: 08/23/2001
Plate: RH.587 row: A column: 2
High quality sequence stop: 527.
Location/Qualifiers
1..644
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RH58702"
/sex="male and female"
/dev_stage="Adult"
/lab_host="DHS-alpha Tona"
/clone_lib="RH Drosophila melanogaster normalized Head
pLC-1"
/note="Organ: head; Vector: pFL1; Site 1: XhoI; Site 2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 232 a 136 c 144 g 132 t
ORIGIN
Query Match 12.7%; Score 325.4; DB 12; Length 644;
Best Local Similarity 99.4%; Pred. No. 3.9e-33;
Matches 337; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GGCCGGCTCCCTTTCGAAAGTCAGTTGGCGGCGGCGGAGATCATCA 60
DB 103 GGCCGGCTCCCTTTCGAAAGTCAGTTGGCGGCGGCGGAGATCATCA 162
QY 61 AACTGGAAGTCGACTCGACTCGAACTGAATTGAACCTGAAAGAAAGAAATATTC 120
DB 163 AACTGGAAGTCGACTCGACTCGAACTGAATTGAACCTGAAAGAAAGAAATATTC 222
QY 121 AAATTGTCGTGTGTGGTGAAGCAAGATATATCTCAAGATATCTGAATACAA 180
DB 223 AAATTGT-GTGTGTGTGGTGAAGCAAGATATATCTCAAGATATCTGAATACAA 281
QY 181 GCTCTGATTTACGAGCAGCAAACTAAGTTACCAATGTGCGAGCGGAAAAAGGAGT 240
DB 282 GCTCTGATTTACGAGCAGCAAACTAAGTTACCAATGTGCGAGCGGAAAAAGGAGT 341
QY 241 GAAAAACGTGCGAATATGCGAGCTAATCTAAAGCATTTGGATTACAAGAAACCGACGAT 300

Db 342 GAAAAAGTGGCAATATGCGAGCTAACTAAAGACATTGGATTACAAAGAACCCAGCAT 401
QY 301 TTGGATTATTAACATTGGACAGGACGAGAAACCTAAG 339
Db 402 TTGGATTATTAACATTGGACAGGACGAGAAACCTAAG 440

RESULT 8
CNS007CM/c 1191 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR15J24 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL067007.1 GI:4945571
VERSION AL067007
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 1191)

REFERENCE
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..1191
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR15J24"
/clone_1lb="RPCI-98"
/note="end : TET3"

BASE COUNT 330 a 254 c 283 g 264 t 60 others
ORIGIN

Query Match 6.1%; Score 155; DB 29; Length 1191;
Best Local Similarity 62.0%; Pred. No. 4,5e-11;
Matches 245; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 720 CTACACAAATGCCCCGACAAATTTGAGGCTAACCCGTTCCGATTCGTCGATGCT 779
Db 503 CCACCTAGTGTATCAAGGATGAGCGCGCCAAAGGAGCCATACCTTCGAGCTCCATGTT 444
QY 780 GGTCTATCGGCGGACCGGCGCTCTGACAGTGTGCGCTTTCTTTAGCTACATGTTCCCTGCA 839
Db 443 CATCTACCGTCCGATTAAGGCTACAGAGTCTGGCGCTTCTCTTCAATGATGATCCCA 384
QY 840 CGCAACTGGTTCCACCTGGGCTTCATATGTCATCAAGCTGTTTGGCATTTCCCT 899
Db 383 TGGCGGATGCTGATCTCGGTTTAAAGTGGAGCTTCAAGTGTGTTTGGGCTGCCACT 324
QY 900 GGAAGTATGACGACGACGCGCAGAGATCGGCGTATCTTACATGCGGCGCTTTTGGCGG 959
Db 323 GGAGATGATTATGATTCCAGAGATAGCTGATCTTATCTTCCGAGGATGCTGGCGG 264

QY 960 ATCCCTGGGACACAGTGTGCTGACTCGAGGTCTTCTGTGGGCGCCAGCGGTGCGT 1019
Db 263 AAGCTGGGACATAGATCTTTGACCCGATGTTCTTTGGGCGCCAGTGTGAGAT 204
QY 1020 CTATGCCCTGTGGCGGACATCTGGCCAACTTACATGAACTATGCGCATGAAG 1079
Db 203 ATACGCTTTGCTAGCCGCCACCTGCAATGCTCTTCAACTATCACCAGATGGTTA 144

QY 1080 CGCATCCAGCAAGCACTGATCGTGTGATCTTTG 1114
Db 143 TGGATCTATCAAGCTCTTCACTTGTGTTTCG 109

RESULT 9
CNS002TC/c 1101 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION BACR07K03 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL063341.1 GI:4941198
VERSION AL063341
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 1101)

REFERENCE
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR07K03"
/clone_1lb="RPCI-98"
/note="end : T7"

BASE COUNT 327 a 253 c 270 g 235 t 16 others
ORIGIN

Query Match 6.0%; Score 154; DB 29; Length 1101;
Best Local Similarity 61.8%; Pred. No. 6,2e-11;
Matches 244; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 720 CTACACAAATGCCCCGACAAATTTGAGGCTAACCCGTTCCGATTCGTCGATGCT 779
Db 505 CCACCTAGTGTATCAAGGATGAGCGCGCCAAAGGAGCCATACCTTCGAGCTCCATGTT 446
QY 780 GGTCTATCGGCGGACCGGCGCTCTGACAGTGTGCGCTTTCTTTAGCTACATGTTCCCTGCA 839
Db 445 CATCTACCGTCCGATTAAGGCTACAGAGTCTGGCGCTTCTCTTCAATGATGATCCCA 386
QY 840 CGCAACTGGTTCCACCTGGGCTTCATATGTCATCAAGCTGTTTGGCATTTCCCT 899

Db	365	TGCCGGGTGCGTGCATCTTCGGTTTAAGTGACGACTAAGCTGGTTTTGGGCTCCCACT	326
Oy	900	GGAGGTGATGCA CGGCAC GCGCAGGANTCGGCCTGATCTAACATGCGCGGGCGTTTTTGCCGG	959
Db	325	GGAATGATGTTCA TGGTTCACAGAGANTAGCTGCATCTTACTTCTCCGGGGGTCTGGCCGG	266
Oy	960	ATCCTCGGGCAC CAGTGTGTGTGATCTCGGAGGTCTTCTCGTGGGGCGCCAGCGTGGCGT	1019
Db	265	AAGCTTGCGCAC TATCATCTTGTGTACCCGGANTGTGTCTCTGTGGCGCCAGTGTGAGGT	206
Oy	1020	CTATCCCCTGTT GGCGCGCACATCTGTGCGCAAATTACATGAACATAAGGCACATAGAAG	1079
Db	205	ATAGCGTTTGTG TAGCCCGCCACCTGGCCAATGTGCTTCTCAANTATCACAGATGCGTTA	146
Oy	1080	CGCATCCACGCA ACTCGGATCGATCGGTGTGATCTTTG	1114
Db	145	TGAGATCATCAAG CTGCTTCATCATCTTGCTTTCG	111
RESULT 10	CNS00E18/c	1101 bp DNA linear GSS 04-JUN-1999	
LOCUS	DROSOPHILA melanogaster genome survey sequence TERT3 end of BAC #		
DEFINITION	BACR29M02 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL069225		
VERSION	AL069225.1 GI:4949368		
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope.		
REFERENCE	Direct Submision		
AUTHORS	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :		
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
COMMENT	determination of this BAC-end and sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.		
FEATURES	location/Qualifiers		
source	1..1101		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7227"		
	/clone="BACR29M02"		
	/clone_1fb="RPCL-98"		
	/note="end : TERT3"		
BASE COUNT	323 a 251 c 242 g 224 t 61 others		
ORIGIN			
Query Match	6.0%; Score 152.4; DB 29; Length 1101;		
Best Local Similarity	63.7%; Pred. No. 1e-10;		
Matches	226; Conservative 3; Mismatches 126; Indels 0; Gaps 0;		
Oy	720	CTACCAATGCCCGCCAGAAATTTGGGCTACCCGTTCCGATCTCGTGGATCGGCTCT	779
Db	498	CCACTCATGTGTGTCAGGGGTGAGCGCGCGCNAAGGGGACCCTATACCTCTGGACTCCATGTT	439

RESULT	11
BF491389	
LOCUS	
DEFINITION	580 bp mRNA linear EST 19-Apr-2001
ACCESSION	AF28124.5prime AT Drosophila melanogaster adult testes POTB7
VERSION	AF28124.5 similar to CG12083:
KEYWORDS	Phen0012083 located on: 3L 62A3-62A3;: 04/09/2001, mRNA sequence.
SOURCE	BF491389.2 GI:13694918
ORGANISM	Drosophila melanogaster (fruit fly)
REFERENCE	Drosophila melanogaster
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
TITLE	1 (bases 1 to 580)
JOURNAL	Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Berman
COMMENT	'B', Carlson, J., Chame, M., Chavez, C., Chew, M., Dorett, V., Farfan
	, P., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P.,
	Liao, G., Miranda, A., Mira, S., Mungall, C. J., Nunoo, J., Pacleb, J.,
	Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E.,
	Celniker, S. and Rubin, G. M.
	BDGP/HMMI AT Drosophila EST Project
	On Dec 6, 2000 this sequence version replaced gi:11574690.
	Contact: Stapleton, M.
	BDGP
	Lawrence Berkeley National Lab
	One Cyclotron Rd, Berkeley, CA 94720, USA
	Fax: 510 486 6798
	Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
	Plate: AT.281 row: B column: 12
	High quality sequence stop: 524.
	Location/Qualifiers
FEATURES	
source	1. 580
	/organism="Drosophila melanogaster"
	/mol_type="mRNA"
	/db_xref="taxon:7227"
	/clone="AT28124"
	/sex="male"
	/dev_stage="0-3 day old Ore-R males"
	/lab_host="Plastes AT.10-AT.120: DH5-alpha. Tona"
	AT.121-AT.319: DH5-alpha. Tona"
	/clone_11b="AT Drosophila melanogaster adult testes POTB7"
	/note="Organ: ADULT testes; Vector: POTB7; Site_1: EcorI;
	Site_2: XhoI; The mRNA for the testis library was made
	from testes and seminal vesicles hand dissected from 0-3
	day old Ore-R males. RNA kindly provided by the lab of
	Margaret Fuller. Sized fractionated cDNAs were directly
	ligated into POTB7. Plasmid cDNA library."
BASE COUNT	101 a 167 c 165 g 147 t

Query Match 5.8%; Score 147.6; DB 10; Length 580;
Best Local Similarity 65.5%; Pred. No. 5.8e-10;
Matches 216; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 720 CTACACAAATGCCCCGAGAAATTTGGGCTAACCCGTTCCGATTCGTCGATTCGGTGT 779
DB 243 CCACTAGTGTATCAGGGGTAGGCGGCCAAGGAGCCCATACCTCCGAGCTCCATGTT 302
QY 780 GGTCTATGCGCGGAGACCGGCGTGCAGTGTGGCGCTTCTTATAGTACATATGTTCTGCA 839
DB 303 CATCTACCTCCGAGTAAAGCTCACAGATCTGGCGCTTCTCTTCAATAGTGTCTCCA 352
QY 840 CGGCAACTGCTTCACCTGGGCTTCAATATGTCATCCAGCTGTTTGGCATTCCTCC 899
DB 363 TGGCGGTGGCTTCATCTCGGCTTTAACTGGCAGTTCAGCTGTTTGGGCTGCCACT 422
QY 900 GGAAGTGTATGACGCGCAGCGCCAGATCGCGTGTATTAATAGCGGGCGTCTTTGGCGG 959
DB 423 GGAATGTGTTCATGTTTCCACGAGATAGCTGATCTCTCTCCGGGTGCTGGCGGG 482
QY 960 ATCCCTGGGACAGATGTGCTGATCTGAGTCTTCTGATGGGCGCCAGCGGTGGCGT 1019
DB 483 AAGCTGGGACCTAGCATCTTTAACCAGATGTGTTCTTGTGGCGCCAGTGTGGAGT 542
QY 1020 CTATGCTGTTGGCGCGCATCTGGCCAA 1049
DB 543 ATACGCTTGTGTCAGCGCCCATCTGGCCAA 572

RESULT 12
CA316643 820 bp mRNA linear EST 26-NOV-2002
LOCUS
DEFINITION UI-M-FMO-cbl-b-05-0-UI.r1 NIH_BMAP_FMO Mus musculus cDNA clone
IMAGE: 6811014 5', mRNA sequence.

ACCESSION CA316643
VERSION CA316643.1 GI:24534767
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@ds-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pyx-5.

FEATURES
Location/Qualifiers
1..820

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6811014"
/issue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH_BMAP_FMO"
/note="Organ: Brain; Vector: pYX-Aac; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Aac vector. The library tag
sequence located between the Not I site and the polyA tail
is AGGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the
Developing Mouse Nervous System", supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

BASE COUNT 145 a 205 c 264 g 204 t 2 others

Query Match 5.3%; Score 135.4; DB 14; Length 820;
Best Local Similarity 53.6%; Pred. No. 1.9e-08;
Matches 305; Conservative 0; Mismatches 261; Indels 3; Gaps 1;

QY 778 CTGCTATATGCGCGGACCGGCGTCTGACAGTGTGGCGCTTCTTTAGTACATGTTCTTG 837
DB 1 CTGCTTACACCCCGCAGCTCCGTCGCGAGGCTTGCGCTATGTGACATCTTCATG 60
QY 838 CAGCCAACTGTTCCACCTGGGCTTCAATATGTCATCCAGCTGTTTGGCATTCCTCC 897
DB 61 CATGCAAGGAGTGAACAGCTGGACTCAATGTGCACTGAGCTGCTGTAGGGGTAATCC 120
QY 898 CTGAGAGTATGACAGCGCAGCGCCAGATCGCGTGTATTAATAGCGGGCGTCTTTGGCG 957
DB 121 CTGGAATGTGTATGAGAGCCAGCCGATTTGGCTTGTCTAGTGGCCGGTGTGTGGCA 180
QY 958 GGAATCCCTGGGACCAAGTGTCTGCACTGGAGGTCTTCTGTTGGGCGCCAGCGGTGGC 1017
DB 181 GGCTCTTGGCGGCTGTCTGTGTGCTGACATGACATGCACTGTTGTGGCTCTTCCGAGGG 240
QY 1018 GTCTATGCGCTTGTGGCGCCGACATCTGGCCACATTAACATGACATATGAGGACATGAAG 1077
DB 241 GTGTATGCGCTGTCTTGTCCCACTAGCAATATGTATGATGATGAGTTCGGGA 300
QY 1078 AGGCAATCAGCAAGCACTCGGATCCGTTGTATCTTGTCTCTCGCATCTGGCTATG-- 1135
DB 301 TGGCAATTCAGATGCTGTGGGATGATGCTGTGCTTGTATGTATGATGAGTTCGGGA 360
QY 1136 -CTCTTACACCCATATCTTCATGATGAGAGCGCTTCCGCAAGGTTCCAGAGTGTCTGAC 1194
DB 361 AGGAGTGTATGCTGCTTGTCCACCATGCGCTTATCCCCCGGCCACCAACCATCTT 420
QY 1195 ATTGCCCACTGACCGGAGCCCTGCGAGACTAAGATGGGCTTGTGATGCTAAGAAC 1254
DB 421 GTGGCACTTGTGGTGGCGGTGGCGGATCACTCGGCGGTGGTCTCAGAAAT 480
QY 1255 TTGCGTACCGAGAGTACGACAGCTCATCTGGTGGCTAGCGTTGGCGCTTACTGTGCC 1314
DB 481 TACGAGACAGAGGTGACGAGACAGTGGCTGTGTGATCTTTGTACATGTACACATC 540
QY 1315 TTACCGCTTTCGCAATCTGTTCAACT 1343
DB 541 TTGCTGTGTTGCTGCTCTTCGAAAT 569

RESULT 13
CA315790 834 bp mRNA linear EST 26-NOV-2002
LOCUS
DEFINITION UI-M-FMO-cdk-g-03-0-UI.r1 NIH_BMAP_FMO Mus musculus cDNA clone
IMAGE: 6810748 5', mRNA sequence.

ACCESSION CA315790
VERSION CA315790.1 GI:24533914
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.ncl.nih.gov/.
1 (bases 1 to 834)

TITLE	JOURNAL	COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished	Contact: Robert Strausberg, Ph.D.

BASE COUNT	149 a	213 c	263 g	206 t	3 others
ORIGIN					

Query Match	5.3%;	Score 135.4;	DB 14;	Length 834;
Best Local Similarity	53.6%;	Pred. No. 1.9e-08;		
Matches 305; Conservative	0;	Mismatches 261;	Indels 3;	Gaps 1

Qy	778	CTGGTCTAATCGCCCGGACCGCGCTGTGACAGGTTGGCCCTCTTTAGCTACAAATGTTCCG	837
Db	25	CTGGTTTAAACACCCGCAAGCTCCGTGGCGAAGGCTTGGCCCTAATGTACAAGTAAATCTTCAATG	84
Qy	838	CACGCCAACTGGTTCACCTTGGGCTTTCAATATCGTATTCACAGCTGTTTTTGGCAATCCC	897
Db	85	CATCCAGGGGTGGAACAGCTGGGACTCATGTGGACATCGCAAGCTCCTGGTAGGGGTATCCC	144
Qy	898	CTGAGAGGTATGCAACGGCAGGCAAGATCGGCGGATCTACATGAGCGGGAGCTTTTGGCC	957
Db	145	CTGAGAGATGTGTCAATGAGGCCAATCCGAATTTGGGCTTTGTCTATGCTGAGCCGGTGTGTGGCA	204
Qy	958	GGATCCCTGGGACACAGTGTGTGTGACATCGGAAGTCTTCCTGGTGGAGCGCAGCGATGAC	1017
Db	205	GGCTCTTTGGGGGTGTGTGTGGCTGTGACATGACTGCACTGTTGTGGGCTCTTCCGAGGGG	264
Qy	1018	GTCATATGCCCTGTGTGGCCGCAATCTGGGCCAAATTACACATGAACATAATGGCGCAATGAAG	1077
Db	265	GTTATATGGGCTGTCTCTGCCCACTTACCCAACTTGTCAATGAATCTGGTCAAGCGATGAAG	324
Qy	1078	AGCGCATCCACGCAACTCGAATCCGTTGTCAATCTTGTCTCCCTCGCATCTGGGCAATG--	1135
Db	325	TGCAGATTCAAGCTGTGGGATGGCTGTGGCTCTTGAATCTGTATAGATATGGAATTTCCGA	384
Qy	1136	-CTTCTTAACCCCAATACTTGTGATGGAAGCGCTTTCGCCAAGGTTCCCAAGGTGTGCTAAC	1194

Db	385	AGGGCTGTATAGCTCCGTTTTCACCCATCGGCTTATCCCCGCTGCCACACCAAGCTTT	444
Qy	1195	ATTGCCCACTGCACGGAGCCCTTGGCAGAGATTACATCGGCTTTCGTGTAAAGAC	125
Db	445	GTGGCACACTTGGGTGGCGTGGCCGTGGGATCACCCCTGGGCGTGGTGTTCAGAAAT	504
Qy	1255	TTGGGTACCGAGAGTAGTAGAGACACTATCTGTGGGCTAGAGGTTGGGGGTCTACTGTGC	1311
Db	505	TACAGACAGAGGCTGCAGAGCAATTCGCTGGTGGAGATCTTTGTGCATGTACACCAATC	564
Qy	1315	TTACACGCTTTGGCCATCGTTTCAACCT	1343
Db	565	TTCTGCTGTGTTGCTGTCTTCTGTGAACAT	593

RESULT 14	
B1544995	
LOCUS	
DEFINITION	B1544995 799 bp mRNA linear EST 05-SEP-2001
ACCSSION	603224246F1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:5284936 5' ,
VERSION	mRNA sequence.
KEYWORDS	B1544995 B1544995
SOURCE	B1544995.1 GI:15432307
ORGANISM	RST .
	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniota; Vertebrate; Euteleostomi;
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 799)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished
	Contact: Robert Strausberg, Ph.D.

BASE COUNT	121 a	270 c	271 g	137 f
ORIGIN				

Query Match	5.1%;	Score 130.8;	DB 12;	Length 799;
Best Local Similarity	61.4%;	Pred. No. 7.6e-08;		
Matches 210; Conservative	0;	Mismatches 132;	Indels 0;	Gaps 0;

Accession	Sequence	Position
Qy	CTGATCATGCGCGACCGGCGTCTCAAGTGTGGGGCTCTTAAAGTACATGTTCCG	837
Db	CTTGTGTACACACCCCGGGACACGTCGCGCGCCCTGGGGCTTCTCACCCTAATGTTATG	455
Qy	CACGCCAATCGTTTCCACTCTGGGCTTCATATGTCATCCAGCTGTTTTTGGACATCC	897

```

DB      456 CACGTTGGGCTGAGAGAGCTGGGGTTCAACGCCCTCTGACGCTGATGATCGGGGTGCCCC 515
QY      898 CTGAGGTGATGACACGGCAGCGACAGGATGAGGATGATCAATGAGCGGCGTTTGGC 957
DB      516 CTGGAATGATGACACGGCCTGCTCGCATAGCCTGCTTCACTGGAGCGGCTGCGCA 575
QY      958 GATATCCCTGGGCAACAGTGTCTGACTCGAGAGTCTTCTGGTGGGCGCGACGCGTGC 1017
DB      576 GGTCTCTTAACCGTCTTCATCAACGACATCGGGCCCCGGTGGTGGAGGCTCGGGGCGG 635
QY      1018 GTCTATGCTCTGTGGCGGACATCTGGCCCACTTAACCTAATGATGCGCAATGAG 1077
DB      636 GTCTAGACCCCTGTGCTCGGACACCTGGCCCAAGTTGTCAATGAACTGGGCTGGAGTGA 695
QY      1078 AGGCATCCACGCAACCTCGGATCCGTTGATCTTTGCTGCC 1119
DB      696 TGTCCTTACAAAGTTGCTGAGAGATGCTGCGCTTGCTGTC 737

RESULT 15
BU454587 712 bp mRNA linear EST 29-NOV-2002
LOCUS     603772383F1 CSEQRBN14 Gallus gallus cDNA clone CHEST707E21 5', mRNA
DEFINITION
ACCESSION BU454587
VERSION    BU454587.1 GI:25943898
KEYWORDS   EST
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
            1 (bases 1 to 712)
REFERENCE  Boardman, P.E., Sanz-Esguero, J., Overton, I.M., Burr, D.W., Bosch, E.,
            Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.,
            A Comprehensive Collection of Chicken cDNAs
            Curr. Biol. 12 (22), 1965-1969 (2002)
            22335534
            PUBMED 12445392
COMMENT    Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology (UMIST)
            )
            PO Box 68, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
            Location/Qualifiers
            1..712
            /organism="Gallus gallus"
            /mol_type="mRNA"
            /strain="Layer"
            /db_xref="taxon:9031"
            /clone="CHEST707E21"
            /sex="Female"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="CSEQRBN14"
            /note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
            EcoRI; Site_2: NotI; This normalized library was
            constructed from 1 million independent clones. cDNA
            synthesis was initiated using an oligo(dT) primer, using
            methylated C in the first strand synthesis reaction.
            Following this first strand reaction, double-stranded cDNA
            was blunt-ended, ligated to NotI adapters, digested with EcoRI
            , size-selected, and cloned into the NotI and EcoRI
            compatible sites of a custom modified MCS of the
            pBluescript (KS+) vector. The library was normalized in 2
            rounds using conditions adapted from Soares et al., PNAS
            (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
            (1996): 791, except that a significantly longer
            reannealing hybridization was used."
FEATURES
source

```

```

BASE COUNT 161 a 182 c 183 g 186 t
ORIGIN
Query Match 4.8%; Score 123.8; DB 13; Length 712;
Best Local Similarity 53.9%; Pred. No. 6,6e-07;
Matches 254; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

QY      647 TCACCGCCAGCACTGGCCCTGGTTCATCTGATGATTCATCATTTAGATTCCATCT 706
DB      169 TACACTGTGCGCTCCGCTGGTTCATGATTAACATCAATCGTAGAGTTGCTTTT 228
QY      707 TCGCTACGACCGCTACACAAATGCGGCCCAAAATTCGGGCTACCGGCTTCGATTCCGT 766
DB      229 TCTTTACATGAGAGTGTCTAGACAAATTTGTACTGCAAGTTACCCCTTATACC 288
QY      767 CGGATTCGGAGCTGTGCTATTCGCGCGGACCGGCGTGTGCAAGTGTGCGGCTTCTTAACT 826
DB      289 TGAATAATGATTAATCTTACCATCTCACTCGTCTGATGAGCTTGAAGTACTAAGCT 348
QY      827 ACATGTTCTGACACGCCAACTGGTTCACCTGGGCTTCAATATCGTCAATCAAGCTGTTT 886
DB      349 ACATATTCATGATGATGAGGAGATAGAACACTTGAAGTCAACGTTGCTTCAAGCTTTGG 408
QY      887 TTGGCATTTCCCTGGAGGTGATGACAGCGACCGCCAGGATCGGCGTGAATCTACATGGCG 946
DB      409 TTGGGCTTCCCTGGAAATGTGTGATGACAGCTGCAAGATCAAGCTTTGTGTATGAGCTG 468
QY      947 GCGTTTTCGCGGATCCCTGGGACCAAGTGTGCTGCACTCGAGAGTCTTCTGTGAGGCG 1006
DB      469 GAGTCTGTGAGAGGCTCTGGGACAGTGTCAATGATGATGATGACGACCTGTGTAGGCT 528
QY      1007 CCAGCGGTGCGTGTATGCTGTGTCGCGCACATCTGGCCCAATTAAGTAACTATG 1066
DB      529 CTTCTGAGAGGTATATGAGCTGTGTCTGAGCTCACTGGCCCAATATATGATGATGATGCT 588
QY      1067 CGCAATGAAGAGCGCATCAGCAACTGGAATCGTTGATCTTTGCT 1117
DB      589 CAGGATGAAGTGCCTCAATCAACTGCTCGCATGCTGTGCTTGATCT 639

```

Search completed: February 8, 2004, 10:22:13
Job time : 6006.41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:35:30 ; Search time 7755.72 Seconds
(without alignments)
10739.422 Million cell updates/sec

Title: US-09-614-150A-11

Perfect score: 2036

Sequence: 1 attcgtctgcgtcgcgtcg.....ttaataactgcacaaattt 2036

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vt:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pla:*
35: em_htg_rtd:*
36: em_htg_man:*
37: em_htg_vtc:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2018.4	99.1	2046	3	AY069405	AY069405 Drosophil
2	1977	97.1	2031	3	DROEMC	M31902 D.melanog
3	1215	59.7	145087	2	AC019753	AC019753 Drosophil
4	1215	59.7	168047	3	AC091219	AC091219 Drosophil
5	1215	59.7	192588	3	AC010057	AC010057 Drosophil
6	1215	59.7	286515	3	AE003469	AE003469 Drosophil
7	1213.4	59.6	1368	3	DROEMC2	M31900 D.melanog
8	1183.4	58.1	1307	3	DROEMC2	M32637 D.melanog
9	822	40.4	934	3	DROEMC1	M32636 D.melanog
10	815.6	40.1	1350	3	DROEMC1	M31901 D.melanog
11	503	24.7	535	3	AB017577	AB017577 Drosophil
12	355.8	17.5	449	3	AF174666	AF174666 Drosophil
13	85.6	4.2	477	5	AF049135	AF049135 Gallus ga
14	85.6	4.2	1239	5	AF068831	AF068831 Gallus ga
15	83.4	4.1	420	6	AX577939	AX577939 Sequence
16	83.4	4.1	420	6	AX577940	AX577940 Sequence
17	83.4	4.1	530	6	AX381227	AX381227 Sequence
18	83.4	4.1	532	6	AX381290	AX381290 Sequence
19	83.4	4.1	533	6	AX381600	AX381600 Sequence
20	83.4	4.1	1049	9	HUMID2HC	D13891 Homo sapien
21	83.4	4.1	1146	10	BC053699	BC053699 Mus muscu
22	83.4	4.1	1276	6	AX305409	AX305409 Sequence
23	83.4	4.1	1276	10	MUSID2A	M69293 Mouse Id2 p
24	83.4	4.1	1280	10	BC006921	BC006921 Mus muscu
25	83.4	4.1	1334	9	BC030639	BC030639 Homo sapi
26	83.4	4.1	2251	10	MMUID2H1	AF077860 Mus muscu
27	83.4	4.1	154126	2	AC010758	AC010758 Homo sapi
28	83.4	4.1	158837	2	AC021659	AC021659 Homo sapi
29	83.4	4.1	177433	9	AC011747	AC011747 Homo sapi
30	83.4	4.1	179243	2	AC116680	AC116680 Mus muscu
31	82.6	4.1	634	6	AX525987	AX525987 Sequence
32	82.6	4.1	725	10	RATID2	D10863 Rattus norv
33	82.6	4.1	268230	2	AC105808	AC105808 Rattus no
34	81.8	4.0	640	6	AX336445	AX336445 Sequence
35	81.8	4.0	640	6	AX578015	AX578015 Sequence
36	81.8	4.0	640	9	HUMID2X	M97796 Human helix
37	80.2	3.9	1172	5	AB019520	AB019520 Xenopus l
38	80.2	3.9	1207	5	BC041527	BC041527 Xenopus l
39	80.2	3.9	1235	5	BT007443	AJ133647 Xenopus l
40	80	3.9	468	9	BT007443	BT007443 Homo sapi
41	80	3.9	468	12	BT008282	BT008282 Synthetic
42	80	3.9	481	6	AX481425	AX481425 Sequence
43	80	3.9	483	6	AX577938	AX577938 Sequence
44	80	3.9	509	6	E07536	E07536 Human id ge
45	80	3.9	509	9	HUMID1HA	D13889 Homo sapien

ALIGNMENTS

RESULT 1
LOCUS AY069405 2046 bp mRNA linear INV 17-DEC-2001
DEFINITION Drosophila melanogaster lD10532 full length cDNA.
ACCESSION AY069405
VERSION AY069405.1 GI:17862145
KEYWORDS FLI_CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 2046)
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
REFERENCE
AUTHORS

Pred. No. is the number of results predicted by chance to have a

TITLE
JOURNAL

Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,
Nunoo, J., Pacle, J., Parag, V., Park, S., Phouanavong, S., Wan, K.,
Yu, C., Lewis, S. E., Rubin, G. M. and Celisner, S.

COMMENT

Submitted (10-DEC-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(http://fruitfly.berkeley.edu) or send email to
cnae@fruitfly.berkeley.edu.

FEATURES

Source

Location/Qualifiers
1..2046
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/map="61D1-61D2"
1..2046
/gene="emc"
/note="Alignment with genomic scaffold AE003469"
/db_xref="FLYBASE:FBgn000575"
243..842
/gene="emc"
/note="Longest ORF"
/codon_start=1
/product="L1D10532p"
/protein_id="AAL39550.1"
/db_xref="GI:17862146"
/db_xref="FLYBASE:FBgn000575"
/translation="MKSILVAVCGTGAGMPLNLSGRILQHPHTRGDNEMKMYLS
KLKDLVPMKPKRLKLTETIIVHIVDITCDIQTELETPENGDALALTAIVGLHE
DESDMEDADAEAEVDPDILAQRLNAEPKAVSPARLPITDRQPTPLVAPHP
QOHQOQOQLOLOQOQOLOSQOQLOLSNSLATPQNAEKDSROS"

gene

CDS

BASE COUNT

726 a 476 c 381 g 463 t

ORIGIN

Query Match 99.1%; Score 2018.4; DB 3; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

17 CTGCAGAAAGATATGAAATGAGAACTTAAACGAGAAACAGCTGGATTAAGT 76
Db 1 CTCGAGAAAGATATGAAATGAGAACTTAAACGAGAAACAGCTGGATTAAGT 60
Qy 77 GTTTCGTTTGAATCGACAGAGTGAAGAAACAAAGTGAAGTGAAGCCGAA 136
Db 61 GTTTCGTTTGAATCGACAGAGTGAAGAAACAAAGTGAAGTGAAGCCGAA 120
Qy 137 ATCAATCAATTAAGAAAGCAGTCAAGAGCAATCTTAACTGAACCTCGAAGTATCC 196
Db 121 ATCAATCAATTAAGAAAGCAGTCAAGAGCAATCTTAACTGAACCTCGAAGTATCC 180
Qy 197 GGTACGACCGGTGAACCAACCGCAAGATTTCTACCAAGAGCGTAAAGCATCATCTCCA 256
Db 181 GGTACGACCGGTGAACCAACCGCAAGATTTCTACCAAGAGCGTAAAGCATCATCTCCA 240
Qy 257 GAATGAAGTCCCTGACGCGCTGTCAGACAGAGTCTCCGAAATGCCGCGCTTGAATG 316

241 GAATGAAGTCCCTGACGCGCTGTCAGACAGAGTCTCCGAAATGCCGCGCTTGAATG 300
Qy 317 CCAGCGGCGCATTCAGACGCGCATCCACGACGCGGCGGAGAAAGCCGAGATGA 376
Db 301 CCAGCGGCGCATTCAGACGCGCATCCACGACGCGGCGGAGAAAGCCGAGATGA 360
Qy 377 AGATGATCTGTCCAACTGAAGAGCTGTGCTTCAATGCCCAAGAACAGAGTCA 436
Db 361 AGATGATCTGTCCAACTGAAGAGCTGTGCTTCAATGCCCAAGAACAGAGTCA 420
Qy 437 CCAAGCTGAGATCATTCAGACAGTCAATGATCAATCTTCCGACCTGACAGCCGAGTGG 496
Db 421 CCAAGCTGAGATCATTCAGACAGTCAATGATCAATCTTCCGACCTGACAGCCGAGTGG 480
Qy 497 AGAGCGACCCCGAGATGGGCACTTGATGTCGCGAGCGCTTGAAGCGGTAAGCGAC 556
Db 481 AGAGCGACCCCGAGATGGGCACTTGATGTCGCGAGCGCTTGAAGCGGTAAGCGAC 540
Qy 557 TCCACGAGACGAGACAGACGATGAGAGATCGGATGCGAGCGAGAAAGCGAAAGTCG 616
Db 541 TCCACGAGACGAGACAGACGATGAGAGATCGGATGCGAGCGAGAAAGCGAAAGTCG 600
Qy 617 ATCCAGATATCTTCGCGCAGCGCTGATATCCGAGACGCGGGAAGTCTTATGTCGG 676
Db 601 ATCCAGATATCTTCGCGCAGCGCTGATATCCGAGACGCGGGAAGTCTTATGTCGG 660
Qy 677 CCGCGCGCTTCGCGCTTACGATGCGCAACGCGCAACCTTGTGAGCGCGCCCATC 736
Db 661 CCGCGCGCTTCGCGCTTACGATGCGCAACGCGCAACCTTGTGAGCGCGCCCATC 720
Qy 737 CGAGACGATCATGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 796
Db 721 CGAGACGATCATGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 780
Qy 797 AGCACTGTGCAACAGTTTACGACGACGACGACGACGACGACGACGACGACGACG 856
Db 781 AGCACTGTGCAACAGTTTACGACGACGACGACGACGACGACGACGACGACGACG 840
Qy 857 AAGGTGGAAGCGCGCGCGGGAAGAAACCGAGGAAACCTTAAAGCTTAAATATG 916
Db 841 AAGGTGGAAGCGCGCGCGGGAAGAAACCGAGGAAACCTTAAAGCTTAAATATG 900
Qy 917 ATATGAAGTATGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 976
Db 901 ATATGAAGTATGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 960
Qy 977 TAACTGTGAAGTCTTCAAGAAACCAAGCCGAGCGCATCTTCAAGCGCTCTTTC 1036
Db 961 TAACTGTGAAGTCTTCAAGAAACCAAGCCGAGCGCATCTTCAAGCGCTCTTTC 1020
Qy 1037 CCGGGAAGCGGAAACCGGGAAGAAAGAGACCTGCTTGGCTTCAAGAGTTAGT 1096
Db 1021 CCGGGAAGCGGAAACCGGGAAGAAAGAGACCTGCTTGGCTTCAAGAGTTAGT 1080
Qy 1097 TAAACATTTAATACAAAGATTAACAACAACAAACCTTACATTAAGAGCGTAAACAA 1156
Db 1081 TAAACATTTAATACAAAGATTAACAACAACAACAAACCTTACATTAAGAGCGTAAACAA 1140
Qy 1157 GTACACATTAATGATTAATTAATTAACCTTCACTAGATGATGCGCTTAAAGATTAAC 1216
Db 1141 GTACACATTAATGATTAATTAATTAACCTTCACTAGATGATGCGCTTAAAGATTAAC 1200
Qy 1217 GGAAGAAAGAAACCTGCGAAGGTTAAACAATTAATTTCTTAAAGCGTTGAAT 1276
Db 1201 GGAAGAAAGAAACCTGCGAAGGTTAAACAATTAATTTCTTAAAGCGTTGAAT 1260
Qy 1277 ACACTGAAGATATATTTTGTCTTAAATTTAGTTTGTGTTTAAACAGCA 1336
Db 1261 ACACTGAAGATATATTTTGTCTTAAATTTAGTTTGTGTTTAAACAGCA 1320
Qy 1337 AGAAAAATGTTACTAATGAAAGCAACAGAGATCGCAATAAAAAATTTAATAA 1396

[illegible][illegible]

Db	781	CAATCGCATATACACGACGAACTGTCCAAACGTTTAGCAAGCCACAGAAATGCGGAGAA	840
Qy	841	GACAGCAGACAGTCGTAAAGCGTGGAAACGCCAGCGGGGAAACCACGAACTTATAG	900
Db	841	GACAGCAGACAGTCGTAAAGCGTGGAAACGCCAGCGGGGAAACCACGAACTTATAG	900
Qy	901	CCTATTAACTAAATGCAATATGAAAGTACGTCAAAATCTCACTCGCAACCTTGTGTGA	960
Db	901	CCTATTAACTAAATGCAATATGAAAGTACGTCAAAATCTCACTCGCAACCTTGTGTGA	960
Qy	961	AGATCTCCGCAATTAAGATAACGATGTAACCTCAAGAAACCAACGCCAGCGCATCTC	1020
Db	961	AGATCTCCGCAATTAAGATAACGATGTAACCTCAAGAAACCAACGCCAGCGCATCTC	1020
Qy	1021	TTCAACGCTCCTTTTCCGGGAAACGAGAAACCCGGGAAACGAGACTCGCTTTTGCT	1080
Db	1021	TTCAACGCTCCTTTTCCGGGAAAGGAGAAACCCGGGAAACGAGACTCGCTTTTGCT	1080
Qy	1081	CCAAACAGAGTTAGTTATTAACATTTATTAACAAGATCAACAACACAAACCTAGCATTT	1140
Db	1081	CCAAACAGAGTTAGTTATTAACATTTATTAACAAGATCAACAACACAAACCTAGCATTT	1140
Qy	1141	AGAGACGGTAAAGCAAGTACATTAATGATTTAATTAACATTAACAAGTACAGATGATG	1200
Db	1141	AGAGACGGTAAAGCAAGTACATTAATGATTTAATTAACATTAACAAGTACAGATGATG	1200
Qy	1201	GCAGTTAAGATGAACGAGAGAGAAAGAAACCTGGCAAGCTTTAAACAATTTAACTTCC	1260
Db	1201	GCAGTTAAGATGAACGAGAGAGAAAGAAACCTGGCAAGCTTTAAACAATTTAACTTCC	1260
Qy	1261	TTTAAACGCTGAATTAACATGAGACATATATTTTGTTCCTATATTTAGTTTGT	1320
Db	1261	TTTAAACGCTGAATTAACATGAGACATATATTTTGTTCCTATATTTAGTTTGT	1320
Qy	1321	TGCTTTAAACAACGGAAGAAAAAGTTACTAATGCAAGCAAGCAAGATCAGAAAT-A	1379
Db	1321	TGCTTTAAACAACGGAAGAAAAAGTTACTAATGCAAGCAAGCAAGATCAGAAAT-A	1380
Qy	1380	AAAAAAATTTAAATAATGCAACGAGCAACCAATCTGAATTTAGTTGATCAACGTTTCT	1439
Db	1381	AAAAAAATTTAAATAATGCAACGAGCAACCAACT---TTACGTTGATCAACGTTTCT	1436
Qy	1440	GCATTAGATTAATGATTTCAATCCAAATGATTTTCTATTAATTTTAATCAATGAAGA	1499
Db	1437	GCATTAGATTAATGATTTCAATCCAAATGATTTTCTATTAATTTTAATCAATGAAGA	1486
Qy	1500	GTTTTCCATTAATGATTAAGCCACAGCCAAATTAACAAGAACAACAATCAGCAGCA	1559
Db	1497	GTTTTCCATTAATGATTAAGCCACAGCCAAATTAACAAGAACAACAATCAGCAGCA	1556
Qy	1560	CACACACATTCAAATTCATTAATTAATTCGAGAGATGAGTTTGAAGTTAAGTTGAG	1619
Db	1557	CACACACATTCAAATTCATTAATTAATTCGAGAGATGAGTTTGAAGTTAAGTTGAG	1616
Qy	1620	AAAGTTCAATTCAGGCACTAATTTTCACTCAACCACTCAATAGAAAAACAACAAC	1679
Db	1617	AAAGTTCAATTCAGGCACTAATTTTCACTCAACCACTCAATAGAAAAACAACAAC	1676
Qy	1680	CTCATTTTGTAGTTTCCGAGTCCCTTGATGCTTTTCTCTTTTAAATGAATTAACG	1739
Db	1677	CTCATTTTGTAGTTTCCGAGTCCCTTGATGCTTTTCTCTTTTAAATGAATTAACG	1736
Qy	1740	CATATGCAATATTAATTAATTCACATTAACCTAAACCTAATAATTAATTAATTA	1799
Db	1737	CATATGCAATATTAATTAATTCACATTAACCTAATAATTAATTAATTAATTA	1786
Qy	1800	TGTTCTGTGATCTGAGCATCTTGTGACAAAGTGTTTTTTTAGCATATTAATGCGATT	1858
Db	1797	TGTTCTGTGATCTGAGCATCTTGTGACAAAGTGT-TTTTTTAGCATATTAATGCGATT	1854
Qy	1860	TTAATTATGATCCGCAATCGAAATTAACAATTTTGTCTCATTTGATGTTAAACAC	1919
Db			

Db	1855	TTAATTATGATCGCAATGAAATTACACATTTTGTCTCAATTCGATTCGATTAAGAACAC	1914
Qy	1920	TCATCATGATGTTAAAAAAGAAAGAAACATTGATGATGTTCCGAAAAATA	1979
Db	1915	TCATCATGATGTTAAAAAAGAAAGAAACATTGATGATGATGTCGAAAAATA	1974
Qy	1980	TATATATAAACATAAAGAAAAACAACCTAGCTCTTAATTAATTTGCAAAATTT	2036
Db	1975	TATATCTTAAACATAAAGAAAAACAACCTAGCTCTTAATTAATTTGCAAAATTT	2031
RESULT 3			
LOCUS	AC019753/c	145087 bp	DNA linear HTG 03-JAN-2000
DEFINITION	Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.		
ACCESSION	AC019753		
VERSION	AC019753.1	GI:6665144	
KEYWORDS	HTG; HTGS PHASE2.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
REFERENCE	Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
AUTHORS	1 (bases 1 to 145087)		
TITLE	Adame, M. and Venter, J.C.		
JOURNAL	Direct Submission		
COMMENT	Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA		
FEATURES	This sequence was identified as CDM:10210835 by the submitter. For more information on this record e-mail to fly@celera.com.		
source	* NOTE: This is a 'working draft' sequence.		
	* This sequence will be replaced		
	* by the finished sequence as soon as it is available and		
	* the accession number will be preserved.		
	Location/Qualifiers		
	1..145087		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7227"		
BASE COUNT	40884 a 32230 c 32629 g 39344 t		
ORIGIN			
Query Match	59.7%; Score 1215; DB 2; Length 145087;		
Best Local Similarity	100.0%; Pred. No. 9, 6e-202; Indels 0; Gaps 0;		
Matches 1215; Conservative	0; Mismatches 0;		
Qy	822	GCCACAGAAATGCGGAGAAAGACAGACAGACAGTGTGAAGGTGGAAACGCCACCGGGAA	881
Db	37659	GCCACAGAAATGCGGAGAAAGACAGACAGACAGTGTGAAGGTGGAAACGCCACCGGGAA	37600
Qy	882	AACCCAGGAACCTTAATAGCTTAACTTAATGAAGTAAGTCAAGTCAAACTCC	941
Db	37599	AACCCAGGAACCTTAATAGCTTAACTTAATGAAGTAAGTCAAGTCAAACTCC	37540
Qy	942	ACTGCAACCTTGCTGGAAGATCTCCGATTAAGTAAGCTGGAACCTCAAGAAAC	1001
Db	37539	ACTGCAACCTTGCTGGAAGATCTCCGATTAAGTAAGCTGGAACCTCAAGAAAC	37480
Qy	1002	CACAGCCACGACGATCTCTTAACGCTCTTTCCCGGAAACGAAAAACCGGGAAA	1061
Db	37479	CACAGCCACGACGATCTCTTAACGCTCTTTCCCGGAAACGAAAAACCGGGAAA	37420
Qy	1062	CGAAGCTCGCTTTGGCTCCAAACAGATTACTTTAAACATTTAATCAAGATACAA	1121
Db	37419	CGAAGCTCGCTTTGGCTCCAAACAGATTACTTTAATTAATTAATTAATTAATTA	37366
Qy	1122	CAACACAAAACCTTAGCATTAAGACAGCGTAAGCAAGTACACATTATGATTTAATTA	1181
Db	37359	CAACACAAAACCTTAGCATTAAGACAGCGTAAGCAAGTACACATTATGATTTAATTA	37300
Qy	1182	CATTACAACTAGATGATGCGCTTAAGATGAAGGAAAGAAACCTTGCAAAAGCT	1241
Db	37299	CATTACAACTAGATGATGCGCTTAAGATGAAGGAAAGAAACCTTGCAAAAGCT	37240

OY		1242	TTAAACAATTAATTTCCTTAAAGCGTGAAAATTACATGGAAGCAATAATTTTTGGTTT	1301
Db		37239	TTAACAATTTAATTTTCCITTAACGCTGAAATTAACCTGAAGCATATATTTTGTTTT	37186
OY		1302	CCTAATTTTAGTTTTGTTTTGGTTTTAAACAACGCAGAAGAAAAATGTTACTAATGCAAACG	1361
Db		37179	CCATATTTTAGTTTTGTTTTGGTTTTAAACAACGCAGAAGAAAAATGTTACTAATGCAAACG	37120
OY		1362	AACAAGATACAGAAATPAAAAAAAATTAATPAAAAATGCCACAGCACCACCAACCTGGAAATTAC	1421
Db		37119	AAACAAGATACAGAAATPAAAAAAAATTAATPAAAAATGCCACAGCACCACCAACCTGGAAATTAC	37060
OY		1422	GTTGAATCACCGTTCCTCGCATTCAGATTAATATCAATCCAAATGATTTTTCCATT	1481
Db		37059	GTTGAATCACCGTTCCTCGCATTCAGATTAATATCAATCCAAATGATTTTTCCATT	37000
OY		1482	AATTTAATCAATGAAAGATTTCCCTAATTAATGATTAAGCAGACCCAAATPAAACAAGAA	1541
Db		36999	AATTTAATCAATGAAAGATTTCCCTAATTAATGATTAAGCAGACCCAAATPAAACAAGAA	36940
OY		1542	ACAACCAACTAGAGAGACACACACATTCCAAATTCATTAATTAATCCAGAGAGATACGT	1601
Db		36939	ACAACCAACTAGAGAGACACACACATTCCAAATTCATTAATTAATCCAGAGAGATACGT	36880
OY		1602	TTAGAGTTTAAGTTCGAGAAAGTTCATTCGAGCATTAATTTTCACTCACACACTCAC	1661
Db		36879	TTAGAGTTTAAGTTCGAGAAAGTTCATTCGAGCATTAATTTTCACTCACACACTCAC	36820
OY		1662	ATGAAAAAACAACAACCTCATTTTGTAGTTTTCCGAGTCCCTTGAGATGTTTTCTCT	1721
Db		36819	ATGAAAAAACAACAACCTCATTTTGTAGTTTTCCGAGTCCCTTGAGATGTTTTCTCT	36760
OY		1722	TTTTTAATCCGAAAPAGCGAATAAGCATTAATTAATTAACCAATATACCTPAAACCTA	1781
Db		36759	TTTTTAATCCGAAAPAGCGAATAAGCATTAATTAATTAACCAATATACCTPAAACCTA	36700
OY		1782	TAAATATATATTAATATATGTTCTGTGATCTGTAGCATCTTTGTCAAAAGTGTTTTT	1841
Db		36699	TAAATATATATTAATATATGTTCTGTGATCTGTAGCATCTTTGTCAAAAGTGTTTTT	36640
OY		1842	TAGCAATAATTATGCGATTTTAAATTATATGATCCGAAATGAAATTAACATTTTTGTTCTCA	1901
Db		36639	TAGCAATAATTATGCGATTTTAAATTATATGATCCGAAATGAAATTAACATTTTTGTTCTCA	36580
OY		1902	TTTGAGATGTPAAAAACACTCATATCATATGATPAAAAAAAACAAAAAGAAACATTGA	1961
Db		36579	TTTGAGATGTPAAAAACACTCATATCATATGATPAAAAAAAACAAAAAGAAACATTGA	36520
OY		1962	TGATGTTGCGAAAAAATATATTAATCTPAAACATPAAAGAAAAACAACACTAGCTCTTAAT	2021
Db		36519	TGATGTTGCGAAAAAATATATTAATCTPAAACATPAAAGAAAAACAACACTAGCTCTTAAT	36460
OY		2022	AAATTGCAAAAATTT 2036	
Db		36459	AAATTGCAAAAATTT 36445	
RESULT 4				
LOCUS	AC091219	168047 bp	DNA	linear INV 09-AUG-2001
DEFINITION	Drosophila melanogaster 3U BAC RP98-4C3 (Roswell Park Cancer			
ACCESSION	AC091219			
VERSION	AC091219.2	GI:15143953		
KEYWORDS	HTG.			
SOURCE	Drosophila melanogaster (fruit fly)			
ORGANISM	Drosophila melanogaster			
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
AUTHORS	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
	Ephydroidea; Drosophilidae; Drosophila.			
	1 (bases 1 to 168047)			
	Muzny,D., Scherer,S., Adams,M.D., Holt,R.A., Evans,C.A.,			

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones.

Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3369-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URI: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

source 1. 168047
location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="3L"
/clone="RP98-4C3"

BASE COUNT 45711 a 37489 c 37230 g 47617 t
ORIGIN

Query Match 59.7%; Score 1215; DB 3; Length 168047;
Best Local Similarity 100.0%; Pred. No. 9,7e-202;
Matches 1215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 822 GCCCAGAAATGCGGAGAAAGCAGCAGACATCGTAAGCGGAAAGCCGCGGAA 881
DB 138074 GCCCAGAAATGCGGAGAAAGCAGCAGACATCGTAAGCGGAAAGCCGCGGAA 138133
QY 882 AACCCAGGAACCTATAGACCTATTAATGATGATGAAGTACAGTCAAACTCC 941
DB 138134 AACCCAGGAACCTATAGACCTATTAATGATGATGAAGTACAGTCAAACTCC 138193
QY 942 ACTCGCAACCTTCTGTGAAGATCTCCGATTAAGATAAGTGAAGTCTCAAGAAAC 1001
DB 138194 ACTCGCAACCTTCTGTGAAGATCTCCGATTAAGATAAGTGAAGTCTCAAGAAAC 138253
QY 1002 CACGAGCCAGGCGATCTCTCAACGCTCTTTCCGCGGAAAGCGGAAACCCGCGGAAA 1061
DB 138254 CACGAGCCAGGCGATCTCTCAACGCTCTTTCCGCGGAAAGCGGAAACCCGCGGAAA 138313
QY 1062 CGAAGACCTCGCTTTGGCTCCAAAGAGTTAGTTAAACATTTAATTAACAAGATACAA 1121
DB 138314 CGAAGACCTCGCTTTGGCTCCAAAGAGTTAGTTAAACATTTAATTAACAAGATACAA 138373
QY 1122 CAACACAAACCTAGCATTTAGAGCGTAAAGCAAGTACACATTATGATTTAATTA 1181
DB 138374 CAACACAAACCTAGCATTTAGAGCGTAAAGCAAGTACACATTATGATTTAATTA 138433
QY 1182 CATTACAACTAGATGATGGCTTAAAGATGAACGAAAGAAAGAAACCTTGCAAACT 1241
DB 138434 CATTACAACTAGATGATGGCTTAAAGATGAACGAAAGAAAGAAACCTTGCAAACT 138493
```

```
QY 1242 TTAACAAATTTAATTTCTTTAAGCCTGAATTAACCTGAAGCATATATTTTGT 1301
DB 138494 TTAACAAATTTAATTTCTTTAAGCCTGAATTAACCTGAAGCATATATTTTGT 138553
QY 1302 CCTAATTTTATTTTGTGTTTGTGTTAAACAGCGCAAGAAATGTTACTAATGAAGC 1361
DB 138554 CCTAATTTTATTTTGTGTTTGTGTTAAACAGCGCAAGAAATGTTACTAATGAAGC 138613
QY 1362 AAACAGATCAGCAATTAATAAATAATTAATAAATGACAGGACCAAACTGAATTA 1421
DB 138614 AAACAGATCAGCAATTAATAAATAATTAATAAATGACAGGACCAAACTGAATTA 138673
QY 1422 GTTGAATCAGCTTTCTCGATTAAGATTAATGATTTTCAATCAATGATTTTCTAT 1481
DB 138674 GTTGAATCAGCTTTCTCGATTAAGATTAATGATTTTCAATCAATGATTTTCTAT 138733
QY 1482 AATTTTATTAATGAAGAGTTTCTTATTAATTAATTAAGCAGCAATTAACAGAGA 1541
DB 138734 AATTTTATTAATGAAGAGTTTCTTATTAATTAATTAAGCAGCAATTAACAGAGA 138793
QY 1542 ACAACCACTCAGCAGACACACACATTCATTAATTTATTAATTCAGAGATGACGT 1601
DB 138794 ACAACCACTCAGCAGACACACACATTCATTAATTTATTAATTCAGAGATGACGT 138853
QY 1602 TTAGAGTTAAGTTGAGAAAGTTCAATTTCCAGGCACTAATTTTCACTCACACACT 1661
DB 138854 TTAGAGTTAAGTTGAGAAAGTTCAATTTCCAGGCACTAATTTTCACTCACACACT 138913
QY 1662 ATGAAAAACAAACCACTCATTTTGTGATTTTCCGAGTCCCTGGATGTTTCTCT 1721
DB 138914 ATGAAAAACAAACCACTCATTTTGTGATTTTCCGAGTCCCTGGATGTTTCTCT 138973
QY 1722 TTTTATTCGAATTAACGCTAATGCTAATTAATTAATTAATTAATTAATTAATTAAT 1781
DB 138974 TTTTATTCGAATTAACGCTAATGCTAATTAATTAATTAATTAATTAATTAATTAAT 139033
QY 1782 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1841
DB 139034 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 139093
QY 1842 TAGCATTAATTAATGATTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1901
DB 139094 TAGCATTAATTAATGATTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 139153
QY 1902 TTTGATGTAAATAACACTCATATCATATGATTAATAAATAAATAAATAAATAAATA 1961
DB 139154 TTTGATGTAAATAACACTCATATCATATGATTAATAAATAAATAAATAAATAAATA 139213
QY 1962 TGATGTTGCAAAATAATATATATATTAATTAATTAATTAATTAATTAATTAAT 2021
DB 139214 TGATGTTGCAAAATAATATATATATTAATTAATTAATTAATTAATTAATTAAT 139273
QY 2022 AATTTGCAAAATTT 2036
DB 139274 AATTTGCAAAATTT 139288
```

RESULT 5
AC010057 192588 bp DNA linear INV 13-JUN-2002
LOCUS Drosophila melanogaster.3L BAC RP98-26C18 (Rowell Park Cancer
DEFINITION Institute Drosophila BAC library) complete sequence.
ACCESSION AC010057
VERSION AC010057.7 GI:21397178
KEYWORDS HMG.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Bukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 192588)
REFERENCE Mazyu, D., Scherer, S., Adams, M.D., Holt, R.A., Evans, C.A.,
Gocayne, J.D., Tabot, P., Williamson, A., Homs, F.H.,

Avella, M., Scott, G.S., Morley, K.W., Amaratunga, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Beeson, K.Y., Brown, M., Butay, C., Busan, D.A., Genter, A., Chen, G., Chen, Z., Clerc-Blankeburg, J., Davisport, L.B., Dietz, S.M., Ding, Y., Dodson, K., Doup, L.E., Draper, H., Emery-Cohen, A., Ferreira, S., Garg, N.D.S., Houck, K., Hostin, D., Howland, T.J., Hume, J., Idegami, C., Jajili, M., Kovar, C., Liu, M., Mattei, B., McIntosh, T.C., Morgan, M., Moy, M., Murphy, B., Nelson, K.A., Nkassas, Y., Nguyen, N., Perez, L., Pitman, G.S., Puri, V., Scheeler, F., Shen, H., Strong, R., Tector, C., Wang, Q., Williams, S.M., Xiang, J., Zaveri, U.S., Zhou, J., Zorrilla, S., Smith, H.O., Wheeler, D., Weinstein, G., Gibbs, R. and Venter, J.C.

TITLE	Direct Submission

REFERENCE

AUTHORS

AUTHORS

2 (bases to 192588)

Worley,K.C., Adams,C., Adio-oduola,B., Ali-osman,F.R., Allen,C., Alebrooks,S.L., Amarrunge,H.C., Are,J.R., Banks,T., Barbara,T., Benton,J., Bimege,K., Blakenburg,K., Bonnah,D., Bouck,U., Bowles,S., Bilewa,M., Brown,E., Brown,M., Bryant,N.P., Bulhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davis,K., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Demu,A.L., Ding,Y., Dinb,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Dublin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,Y., Garcia,A., Garner,T., Garza,N., Gill,R., Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hernandez,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsli,F., Howard,S., Huber,J., Hulyk,S., Hume,U., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudash,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisleged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,Y., Maheshwari,M., Mapa,P., Martin,R., Matindale,A., Martinez,E., Massey,E., Mawhinney,E., Mcleod,M.P., Meador,M., Mel,G., Meeker,M., Miner,G., Miner,G., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogunu,M., Okunodu,G., Oragunye,N., Oviedo,R., Face,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojchoudhary,I., Rojfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostlari,N., Stinson,I., Sodergren,E., Sonatke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,Y., Villallon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Washington,S., Williams,G., Williamson,A., Wlaczek,R., Wooden,S., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

TITLE Direct Submission

JOURNAL REFERENCE

AUTHORS

AUTHORS

Direct Submission
Submitted (11-SEP-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

AUTHORS

AUTHORS

Albrooks, S. L., Ameratunga, H. C., Are, J. R., Banke, T., Barbata, J., Benson, J., Bimege, K., Blankenburg, K., Bonin, D., Bouck, J., Bowie, S., Brielev, M., Brown, E., Brown, M., Bryant, N. P., Buhay, C., Burck, P., Burkett, C., Butrell, K. L., Byrd, N. C., Carron, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dethorne, S. R., David, R., Davila, M. L., Davis, C., Dean, A. L., Ding, Y., Dinh, H. H., Donthwaite, K. J., Draper, H.,

REFERENCE

AUTHORS

TITL

JOURNAL

COMMENT

Dugan-Roch S, Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gotrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J., Hernandez, O., Hodson, A., Hogue, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Huylk, S., Hume, Y., Joudah, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Koryah, J., Kovar, C., Kratovic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Lounsbeg, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, Y., Maheshwari, M., Mapa, P., Martin, R., Matindale, A., Martinez, E., Massey, E., Mawhinney, E., Meleod, M.P., Mesdor, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, N., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkenkwo, S., Oguh, M., Okunnu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoochattai, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabot, P., Tamelison, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usumi, K., Vasquez, L., Vera, V., Villation, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission
Submitted (13-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
(bases 1 to 192588)

TITLE

JOURNAL

REFERENCE

AUTHORS

TITL

Submitted (13-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 13, 2002 this sequence version replaced gi:15594190.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STRS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING LEAD COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base

TITLE
JOURNAL
MEDLINE
PUBMED
AUTHORS
REFERENCE
2
(bases 1 to 286515)
2
10731132
Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C.,
Stapleton, M., Strong, R., Sun, X., Svirska, R., Tector, C., Turner, R.,
Venster, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A.,
Weinstock, G.M., Weisenbach, J., Williams, S.M., Woodrager,
Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F.,
Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H.,
Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O.,
Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
The genome sequence of *Drosophila melanogaster*
Science 287 (5461), 2185-2195 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
AUTHORS
REFERENCE
3
(bases 1 to 286515)
3
10731132
Mitra, S., Crosby, M.A., Matthews, B.B., Bayraktaroglu, L.,
Campe, S., Hradecky, P., Huang, Y., Kaminker, J.S., Prochnik, S.E.,
Smith, C.D., Tupy, J.L., Bergman, C.M., Berman, B.P., Carlson, J.W.,
Clinker, S.E., Clump, M.E., Drysdale, R.A., Emmert, D., Frise, E., de
Grey, A.D.N.J., Harris, N.L., Kromm, B., Marshall, B.,
Millburn, G.H., Richter, J., Russo, S., Searle, S.W.J., Smith, E.,
Shu, S., Smutniak, F., Whitfield, B.J., Ashburner, M., Gelbart, W.M.,
Rubin, G.M., Mungall, C.J. and Lewis, S.E.
Annotation of *Drosophila melanogaster* genome
Unpublished

TITLE
JOURNAL
MEDLINE
PUBMED
AUTHORS
REFERENCE
4
(bases 1 to 286515)
4
10731132
Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
5
(bases 1 to 286515)

TITLE
JOURNAL
MEDLINE
PUBMED
AUTHORS
REFERENCE
5
(bases 1 to 286515)
5
10731132
FlyBase
Direct Submission
Submitted (06-SEP-2002) University of California Berkeley, 539 Life
Sciences Addition, Berkeley, CA 94720, USA
6
(bases 1 to 286515)

TITLE
JOURNAL
MEDLINE
PUBMED
AUTHORS
REFERENCE
6
(bases 1 to 286515)
6
10731132
FlyBase
Direct Submission
Submitted (23-JUN-2003) University of California Berkeley, 539 Life
Sciences Addition, Berkeley, CA 94720, USA
On Sep 17, 2002 this sequence version replaced gi:10727202.
Location/Qualifiers
1. 286515
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="3L"
/note="genotype: y[1], cn[1] bw[1] ep[1], Rh6[1]"
20999. .23795
/locus_tag="CG12030"
/note="last curated on Thu May 16 11:05:06 PDT 2002"
/map="61C8-61C8"
/db_xref="FLYBASE:FBgn0035147"
/join(20999. .21533,21661. .21901,22715. .23005,23066. .23795)
/locus_tag="CG12030"
/product="CG12030-RA"
/db_xref="FLYBASE:FBgn0035147"

CDS
join(21662. .21901,22715. .23005,23066. .23587)
/locus_tag="CG12030"
/note="CG12030 gene product"
/codon_start=1
/product="CG12030-PA"
/protein_id="AA047398.1"
/db_xref="GI:7291983"
/db_xref="FLYBASE:FBgn0035147"
/translation="MAPTVLTGAGVYSGSHVLEMLNAGVVICVDNLCAVSSGA
KLPALSRQELTGKRVNRYDITDREQVRSYFQHKIDMAVHPALAAVSGSCIP
LQYHNMTGTVLLEAMADNVKFTVSSATVGEKFLPTVTEHPGNCSPYK
TKYTEELIKDLCKSDKRWAVSLRYFNVGAIISGRIGEDNPGENMLAPYLAQAV
GRPSLSVIGSDPFTDGTGVVDYIHLVLAEGHVALLDKILNIATGFPAYVLGTGV
GYSVLNVAFKASGSKVNYTLVDRSGDVATCVADATLADKKLGWKAERGIDKNCE
DTRWMSQNPNGANK"
complement(24020. .25316)
/locus_tag="CG3402"
/note="last curated on Thu May 16 08:49:06 PDT 2002"
/map="61C8-61C8"
/db_xref="FLYBASE:FBgn0035148"
complement(join(24020. .24758,24926. .25054,25193. .25316))
/locus_tag="CG3402"
/product="CG3402-PA"
/db_xref="FLYBASE:FBgn0035148"
complement(join(24570. .24758,24926. .25054,25193. .25246))
/locus_tag="CG3402"
/note="CG3402 gene product"
/codon_start=1
/product="CG3402-PA"
/protein_id="AA047399.1"
/db_xref="GI:7291984"
/db_xref="FLYBASE:FBgn0035148"
/translation="MAKMAFOHQAAGTAMCISPTLHAKKDYDQSGREILKGFPG
GGIDQYKKSPOGYTDYGIYTVFHEGSPALRAGRIHDKILLQCNQDTPTWTHKAV
STIRKPLNLMVARKGVST"
complement(25463. .26680)
/gene="Trap25"
/locus_tag="CG17183"
/note="last curated on Thu May 16 08:48:39 PDT 2002"
/map="61C8-61C8"
/db_xref="FLYBASE:FBgn0035149"
complement(25463. .26680)
/gene="Trap25"
/locus_tag="CG17183"
/product="CG17183-RA"
/db_xref="FLYBASE:FBgn0035149"
complement(25536. .26492)
/gene="Trap25"
/locus_tag="CG17183"
/note="Trap25 gene product"
/codon_start=1
/product="CG17183-PA"
/protein_id="AA047400.1"
/db_xref="GI:7291985"
/db_xref="FLYBASE:FBgn0035149"
/translation="MKKYGQNGQNGQPSGGGGGPGNNMPKGRGMOHGNQOMMS
PQHQQQQQMGMGGPGSMQMPNPGFPGGGLAPGSPQHQQQQQOMQOMQMP
QGGVGVGMGGGVGGGQVPPQQQQQPMQPPQPPQQQQQINPVAGIPPGAG
GSNNMLAIQQNPHEINIVOLSRLOETVODIASRFQVFAKLKQIPTSHRENSB
KRYQEFRTIRLLFKRVRIIYKCDNAGDYNASAEILYRDPPEPRIRIPSCDEYRK
VLOENHELETVALKRQREIIRIIMWEINTMLMARRS"
26843. .30457
/gene="Rev1"
/locus_tag="CG12189"
/note="last curated on Wed May 15 08:10:08 PDT 2002"
/map="61C8-61C8"
/db_xref="FLYBASE:FBgn0035150"
/join(26843. .26959,27018. .28375,28430. .28712,28775. .29198,
29261. .29314,29380. .30067,30132. .30457)
/gene="Rev1"
/locus_tag="CG12189"
/product="CG12189-RA"
/db_xref="FLYBASE:FBgn0035150"

CDS

join(26924, .26959,27018, .28375,28430, .28712,28775, .29198, 29261, .29314,29380, .30067,30132, .30276)
 /gene="Rev1"
 /locus_tag="CG12189"
 /note="Rev1 gene product"
 /codon_start=1
 /product="CG12189-PA"
 /protein_id="AAFA7401.1"
 /db_xref="GI:7291986"
 /db_xref="FLYBASE:FBgn0035150"
 /translation="MTREDNDNGSEKSGYFEAKSKYLEBPAAASDPFRKSDLFGGIS IFVNGRTPSDAEIKRIMVNGSTFFHHERSHYTYIIASVLPDKVKNMNLSTKTSK MWVDCLEKKKIVDPKPYLLYTNKTSQPMILFGKPKNGNANESSDVEPKDAVEV

Query Match 59.7%; Score 1215; DB 3; Length 286515;
 Best Local Similarity 100.0%; Pred. No. 1e-201;
 Matches 1215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 822 GCCACAGATGCGAGAAAGACAGACAGACAGTCTGTAAGCGTGGAAACGCCGAGCGGGA 881
 124753 GCCACAGATGCGAGAAAGACAGACAGACAGTCTGTAAGCGTGGAAACGCCGAGCGGGA 124812

QY 882 AACCCAGGAACTTATTAAGCTTATTAATGATATGATAAGAGTACAGTCAAACTCC 941
 124813 AACCCAGGAACTTATTAAGCTTATTAATGATATGATAAGAGTACAGTCAAACTCC 124872

QY 942 ACTGCAACCTTCTGTAAGATCTCCGATTAAGTATGATGTAAGTCAAGAACTCAAGAAAC 1001
 124873 ACTGCAACCTTCTGTAAGATCTCCGATTAAGTATGATGTAAGTCAAGAACTCAAGAAAC 124932

QY 1002 CACAGCCGAGCGGATCTCTTCAACGCTCTTTCCCGGAAACGGAAACCGCGGAAA 1061
 124933 CACAGCCGAGCGGATCTCTTCAACGCTCTTTCCCGGAAACGGAAACCGCGGAAA 124992

QY 1062 CGAGACTCGCTTTGGCTCCACAGAGTTTAACTTTAACTTTATTAACAAGATACAA 1121
 124993 CGAGACTCGCTTTGGCTCCACAGAGTTTAACTTTAACTTTATTAACAAGATACAA 125052

QY 1122 CAACCAAAACCTTGATTAAGAGCGGTAAGCAAGTACATTAATGATTTATTAATAA 1181
 125053 CAACCAAAACCTTGATTAAGAGCGGTAAGCAAGTACATTAATGATTTATTAATAA 125112

QY 1182 CATTACAAGTATGATGAGCGCTTAAGATGAACGAAAGAAAGAACTTGCGAAAGCT 1241
 125113 CATTACAAGTATGATGAGCGCTTAAGATGAACGAAAGAAAGAACTTGCGAAAGCT 125172

QY 1242 TTAAACAATTTATTTCTTTTAAAGCTTGAATTAACCTGAGCATATTTTGTCTT 1301
 125173 TTAAACAATTTATTTCTTTTAAAGCTTGAATTAACCTGAGCATATTTTGTCTT 125232

QY 1302 CCTAATTTAGTTTGTGTTTAAACAAGCAAGAAAGAAAGTCTTAATGCAAGC 1361
 125233 CCTAATTTAGTTTGTGTTTAAACAAGCAAGAAAGAAAGTCTTAATGCAAGC 125292

QY 1362 AAACGAGATGAGCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 1421
 125293 AAACGAGATGAGCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 125352

QY 1422 GTTGAATCAGCGTTCTCTGATTAAGTATTAATTAATTAATTAATTAATTAATTAAT 1481
 125353 GTTGAATCAGCGTTCTCTGATTAAGTATTAATTAATTAATTAATTAATTAATTAAT 125412

QY 1482 AATTTATTAATGAAGAGTTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1541
 125413 AATTTATTAATGAAGAGTTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 125472

QY 1542 ACAACCACTGAGCAGACACACACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1601
 125473 ACAACCACTGAGCAGACACACACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 125532

QY 1602 TTAGAGTTTAAGTTGAGAAAGTTTCAAGGCACTAATTTTCACTCAACCACTCAC 1661
 125533 TTAGAGTTTAAGTTGAGAAAGTTTCAAGGCACTAATTTTCACTCAACCACTCAC 125592

QY 1662 ATGAAAAACAAACACCTCATTTTGTAGTTTTCGAGTCCCTTGATGTTTCTCT 1721
 125593 ATGAAAAACAAACACCTCATTTTGTAGTTTTCGAGTCCCTTGATGTTTCTCT 125652

QY 1722 TTTTATTCGAATTAACGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1781
 125653 TTTTATTCGAATTAACGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 125712

QY 1782 TAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1841
 125713 TAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 125772

QY 1842 TAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1901
 125773 TAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 125832

QY 1902 TTTTGATGTTAAACACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1961
 125833 TTTTGATGTTAAACACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 125892

QY 1962 TGAATTTGCGAAAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2021
 125893 TGAATTTGCGAAAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 125952

QY 2022 AAATTCGAAAAATTT 2036
 125953 AAATTCGAAAAATTT 125967

RESULT 7
 DROEMC2 1368 bp DNA linear INV 26-APR-1993
 LOCUS D.melanogaster extramicrochaetae (emc) protein gene, exon 2.
 ACCESSION M31900
 VERSION M31900.1 GI:157348
 KEYWORDS extramicrochaetae protein.
 SEGMENT 2 of 2
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1368)
 Garreil,J., and Modolell,J.
 JOURNAL Unpublished (1990)
 AUTHORS 2 (bases 84 to 1368)
 Garreil,J., and Modolell,J.
 TITLE The Drosophila extramicrochaetae locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein
 JOURNAL Cell 61 (1), 39-48 (1990)
 MEDLINE 90199895
 PUBMED 1690605
 COMMENT Original source text: D.melanogaster (strain Canton S) DNA, clones lambda EMC [3,7].
 Draft entry and computer-readable sequence for [2] kindly submitted by J.Garreil, 05-FEB-1990.
 FEATURES
 source location/Qualifiers
 1..1368
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 join(M31901.1:635..1198,84..119)
 /note="extramicrochaetae protein (emc)"
 /codon_start=1
 /protein_id="AA28510.1"
 /db_xref="GI:157350"
 /translation="MKSILAVCOTGASGMPALNASGRIQHPTRHGENVAEMKMYLS KLDIVFPMKRNKLTLEIIOHIVYICDLOTELETHPEMGFDAALATVANGLHE DEBSDEPDABEAEVDPDVLQRIANRBPATVSSPAALPLTDGTPTVLAHPH QQH0Q0Q0LOLQ0Q0LO0Q0Q0LSNLSLTPPNAEKDSR0S"
 prim_transcript <1..1297
 /note="emc mRNA and introns"

intron <1..83
 /note="emc intron A"
 exon 84..>119
 /note="extramacrochaetae protein (emc)"
 /number=2
 BASE COUNT 520 a 260 c 196 g 392 t
 ORIGIN About 2.5 kb after segment 1.
 Query Match 59.6%; Score 1213.4; DB 3; Length 1368;
 Best Local Similarity 99.9%; Pred. No. 1.1e-201;
 Matches 1214; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 822 GCCACAGATGCGGAGAAAGACAGACAGACAGTCTGTAAGCGTGAAGCGCCAGCGGAAA 881
 DB 83 GCCACAGATGCGGAGAAAGACAGACAGACAGTCTGTAAGCGTGAAGCGCCAGCGGAAA 142
 QY 882 AACCCAGGAACCTATTAAGCTTAACTAATGATATGAAGTACAGTCAAACTCC 941
 DB 143 AACCCAGGAACCTATTAAGCTTAACTAATGATATGAAGTACAGTCAAACTCC 202
 QY 942 ACTGCAACCTTGGTGAAGATCTCGCATATGATTAAGTGAAGCTGCAAGAAAC 1001
 DB 203 ACTGCAACCTTGGTGAAGATCTCGCATATGATTAAGTGAAGCTGCAAGAAAC 262
 QY 1002 CACAGCCAGCGCATCTTCAACGCTCTTTCCCGGAAAACGGAACCCGGGAAA 1061
 DB 263 CACAGCCAGCGCATCTTCAACGCTCTTTCCCGGAAAACGGAACCCGGGAAA 322
 QY 1062 CGAGACCTCGCTTTGGCTCCACAGAGTTTAACTTTAATTAACAGAAATCAA 1121
 DB 323 CGAGACCTCGCTTTGGCTCCACAGAGTTTAACTTTAATTAACAGAAATCAA 382
 QY 1122 CAACAGAAACCTTGAAGAGAGGTAAGAGATGACATTAATGATTAATTAATTA 1181
 DB 383 CAACAGAAACCTTGAAGAGAGGTAAGAGATGACATTAATGATTAATTAATTA 442
 QY 1182 CATTACAAGTATGATGAGCGCTTAAGATGAACGAGAAAGAAACTTGGCAAGCT 1241
 DB 443 CATTACAAGTATGATGAGCGCTTAAGATGAACGAGAAAGAAACTTGGCAAGCT 502
 QY 1242 TTAACAATTTAATTTCTTTAAAGCTTGAATTAACCTGAACATATATTTTGT 1301
 DB 503 TTAACAATTTAATTTCTTTAAAGCTTGAATTAACCTGAACATATATTTTGT 562
 QY 1302 CCTAATTTAGTTTGTGTTTAAACGAGAAAGAAAGTATCTATATGCAAGC 1361
 DB 563 CCTAATTTAGTTTGTGTTTAAACGAGAAAGAAAGTATCTATATGCAAGC 622
 QY 1362 AAACAGAGATCAGCAATTAATAAATTAATAAATGACAGGCAACCACTGGAATTAC 1421
 DB 623 AAACAGAGATCAGCAATTAATAAATTAATAAATGACAGGCAACCACTGGAATTAC 682
 QY 1422 GTTGAATCAGCGTTCTGATTAAGATTAATGATTAATCAATCAAAATGATTTCT 1481
 DB 683 GTTGAATCAGCGTTCTGATTAAGATTAATGATTAATCAATCAAAATGATTTCT 742
 QY 1482 AATTTAATCAATGAAGATTTCTATATATGATTAAGCAAGCAATTAACCAAGA 1541
 DB 743 AATTTAATCAATGAAGATTTCTATATATGATTAAGCAAGCAATTAACCAAGA 802
 QY 1542 ACAACCAACTCAGAGACACACATTCAAATCATTAATTTATCGAGAGTACGT 1601
 DB 803 ACAACCAACTCAGAGACACACATTCAAATCATTAATTTATCGAGAGTACGT 862
 QY 1602 TTAGAGTTAAGTTGAGAAAGTTCATTTCCAGGACATTAATTTCACTCAACCACTCAC 1661
 DB 863 TTAGAGTTAAGTTGAGAAAGTTCATTTCCAGGACATTAATTTCACTCAACCACTCAC 922
 QY 1662 ATGAGAAAAACAAACCTCATTTTGTAGTTTCCAGATCCCTTGATGTTTCTCT 1721
 DB 923 ATGAGAAAAACAAACCTCATTTTGTAGTTTCCAGATCCCTTGATGTTTCTCT 982
 QY 1722 TTTTATCGAAATTAAGCATATATGATTAATTAATTAATTAATTAATTAATTA 1781

DB 983 TTTTATCGAAATTAAGCATATGATTAATTAATTAATTAATTAATTAATTAATTA 1042
 QY 1782 TAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1841
 DB 1043 TAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1102
 QY 1842 TAGCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1901
 DB 1103 TAGCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1162
 QY 1902 TTTCGATGTAATAACACTTATCATGATGTAATAATAATAATAATAATAATAATAATA 1961
 DB 1163 TTTCGATGTAATAACACTTATCATGATGTAATAATAATAATAATAATAATAATAATA 1222
 QY 1962 TGATGTTGCGAAAAATATTTATCTTAATAACATTAAGAAAAACAAACCTAGCTTTAAT 2021
 DB 1223 TGATGTTGCGAAAAATATTTATCTTAATAACATTAAGAAAAACAAACCTAGCTTTAAT 1282
 QY 2022 AAATGCAAAAAATTT 2036
 DB 1283 AAATGCAAAAAATTT 1297
 RESULT 8
 DROEMCA2
 LOCUS
 DEFINITION D.melanogaster extramacrochaetae protein gene, exon 2.
 ACCESSION M32637.1 GI:157352
 VERSION
 KEYWORDS extramacrochaetae protein; helix-loop-helix protein; regulatory protein.
 SEGMENT
 SOURCE 2 of 2.
 ORGANISM Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE
 AUTHORS Ellis,H.M., Spann,D.R. and Posakony,J.W.
 TITLE 1 (bases 1 to 1307)
 JOURNAL Cell 61 (1), 27-38 (1990)
 MEDLINE 90199894
 PUBMED 1690604
 COMMENT Original source text: D.melanogaster 8-12 hour embryo, cDNA to mRNA, clones pMB5, and DNA, clones pF7, lambda-B[11,16]. Draft entry and computer-readable sequence for [Cell (1990) In press] kindly submitted by J.W.Posakony, 06-MAR-1990.
 FEATURES
 source Location/Qualifiers
 1..1307
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 join(M32636.1:347..910,31..66)
 /note="extramacrochaetae protein (emc)"
 /codon_start=1
 /protein_id="AA28511.1"
 /db_xref="GI:157354"
 /translation="MKSILRAYCQASGMPALNAGRIORHPTHRGDGENAEMKYLS KLDIVPEPKRKRKLTKEITIGHVIDYCDLTETETHEKGNFPAALATVNGIHE DEDSDMEDADAEAEVDPDILAQRLNAEQPAKVSSPARLP/LTRQNTNTLVAPAP OHOOQOQLOLOOQOLOSQOOLSNSLATPQNAEKSRQS"
 prim_transcript <1..1140
 /note="emc mRNA and introns"
 intron <1..30
 /note="emc intron"
 exon 31..>66
 /note="extramacrochaetae protein (emc)"
 /number=1
 BASE COUNT 499 a 244 c 193 g 371 t

Db 209 AGGCGACAAAGCCGAAATCAATTAATGAAAAAGCCAGTCAGAGCAATCTTTAACTGA 268
Qy 181 ACCTCGCAAGTAATCCGGTACGACCGGTGTAAACAACCGCAAGATTCTACTCAGAGAGCGT 240
Db 269 ACCTCGCAAGTAATCCGGTACGACCGGTGTAAACAACCGCAAGATTCTACTCAGAGAGCGT 328
Qy 241 AAAGCAATCATCTCCAGAAATGAAGTCCCTGACGCGCTCTGCGACAGAGTGCCTCCGGA 300
Db 329 AAAGCAATCATCTCCAGAAATGAAGTCCCTGACGCGCTCTGCGACAGAGTGCCTCCGGA 388
Qy 301 ATCCCGCTTTGAATGCGACGCGCGCATCCAGCGCATCCACGACCGCGCGACGCGG 360
Db 389 ATCCCGCTTTGAATGCGACGCGCGCATCCAGCGCATCCACGACCGCGCGACGCGG 448
Qy 361 GAGAACGCCGAGATGAAGATGTATCTGTCCAAACTGAAGAGCTCGTTCGTTCAATGCC 420
Db 449 GAGAACGCCGAGATGAAGATGTATCTGTCCAAACTGAAGAGCTCGTTCGTTCAATGCC 508
Qy 421 AAGAACGAGAAAGCTCAACCAAGCTGAGATCATCCAGACGTCATCGACTCATCTGCGAC 480
Db 509 AAGAACGAGAAAGCTCAACCAAGCTGAGATCATCCAGACGTCATCGACTCATCTGCGAC 568
Qy 481 CTGCAAGACGAGCTGAGACGACACCCGAGATGGCACTTCATGCGCGACGCGCTCTG 540
Db 569 CTGCAAGACGAGCTGAGACGACACCCGAGATGGCACTTCATGCGCGACGCGCTCTG 628
Qy 541 AGCGCGGTGAAGGAGCTCCACGAGACGAGACGAGACGAGATGGAGATGGCGATGCCGAG 600
Db 629 AGCGCGGTGAAGGAGCTCCACGAGACGAGACGAGACGAGATGGAGATGGCGATGCCGAG 668
Qy 601 GCAGAAAGCGAAAGTCCATCCAGATATCTCGCCAGCGCTGATGCGCGACGCGAGCGG 660
Db 689 GCAGAAAGCGAAAGTCCATCCAGATATCTCGCCAGCGCTGATGCGCGACGCGCGG 748
Qy 661 AAAGTCTTAGTCCCGCGCGCTCTCCGCTTACCGATGCCAAAGCGCCAACTCTTT 720
Db 749 AAAGTCTTAGTCCCGCGCGCTCTCCGCTTACCGATGCCAAAGCGCCAACTCTTT 808
Qy 721 GTGGCGCGCGCGCATCCGAGACGAGATCCAGACGAGACGAGATCCGAACTGCGACGCA 780
Db 809 GTGGCGCGCGCGCATCCGAGACGAGATCCAGACGAGACGAGATCCGAACTGCGACGCA 868
Qy 781 CAAGTCAATCAGACGACGAACTGTCCACAGTTTGAACAG 822
Db 869 CAAGTCAATCAGACGACGAACTGTCCACAGTTTGAACAG 910

RESULT 10
DROEMC1 1350 bp DNA linear INV 26-Apr-1993
LOCUS D.melanogaster extramicrochaetae (emc) protein gene, exon 1.
ACCESSION M31901.1 GI:157347
VERSION M31901.1 GI:157347
KEYWORDS extramicrochaetae protein.
SEGMENT 1 of 2
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1350)
Garell,J. and Modolell,J.
REFERENCE JOURNAL Unpublished (1990)
AUTHORS 2 (bases 1 to 1198)
REFERENCE Garell,J. and Modolell,J.
TITLES The Drosophila extramicrochaetae locus, an antagonist of proneural
genes that, like these genes, encodes a helix-loop-helix protein
JOURNAL Cell 61 (1), 39-48 (1990)
MEDLINE 90198985
PUBMED 1690605
Original source text: D.melanogaster (strain Canton S) DNA, clones
lambda EMC[3,7].
Draft entry and computer-readable sequence for [2] kindly submitted

by J. Garell, 05-FEB-1990.
FEATURES
Location/Qualifiers
source 1..1350
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
prim_transcript 377..>1350
/note="emc mRNA and introns"
exon <635..1198
/note="extramicrochaetae protein (emc)"
intron 1199..>1350
/note="emc intron A"
BASE COUNT 400 a 375 c 316 g 259 t
ORIGIN Chromosome 3L.
Query Match 40.1%; Score 815.6; DB 3; Length 1350;
Best Local Similarity 99.5%; Pred. No. 36-132;
Matches 818; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATTGCTGTCTGACCGCTGCGCAAAAGATATGAAATCGAAACGTAAAGCGAAGAA 60
Db 377 ATTGCTGTCTGACCGCTGCGCAAAAGATATGAAATCGAAACGTAAAGCGAAGAA 436
Qy 61 CAGCTGATTAAGTGTTCGTTTGAACATCGACAGAGTGAAGAAACAAAGTGAAGT 120
Db 437 CAGCTGATTAAGTGTTCGTTTGAACATCGACAGAGTGAAGAAACAAAGTGAAGT 496
Qy 121 AAGCGCAAAAGCGAAATCAATCAATTAAGAAAGCACTCAAGCAATCTTTAACTGA 180
Db 497 AAGCGCAAAAGCGAAATCAATCAATTAAGAAAGCACTCAAGCAATCTTTAACTGA 556
Qy 181 ACTCGCAAGTATCCGCTGACACCGGTGAACAACCGCCAAATTTACTCAAGAGCGT 240
Db 557 ACTCGCAAGTATCCGCTGACACCGGTGAACAACCGCCAAATTTACTCAAGAGCGT 616
Qy 241 AAAGCAATCATCTCCAGATGAAGTCCCTGACGCGCGCTGCGACAGAGTCCCGGA 300
Db 617 AAAGCAATCATCTCCAGATGAAGTCCCTGACGCGCGCTGCGACAGAGTCCCGGA 676
Qy 301 ATGCGCGCTTGAATGCGACGCGCGCATCCAGCGCATCCACGACCGCGCGCAAGCGG 360
Db 677 ATGCGCGCTTGAATGCGACGCGCGCATCCAGCGCATCCACGACCGCGCGCAAGCGG 736
Qy 361 GAGAACGCCGAGATGAAGTATCTGTCCAACTGAAAGACCTGTTCCGTTCAATGCC 420
Db 737 GAGAACGCCGAGATGAAGTATCTGTCCAACTGAAAGACCTGTTCCGTTCAATGCC 796
Qy 421 AAGAACGAGAGCTCAACCAAGCTGAGATCATCCAGACGTCATGACTACATCTGCGAC 480
Db 797 AAGAACGAGAGCTCAACCAAGCTGAGATCATCCAGACGTCATGACTACATCTGCGAC 856
Qy 481 CTGCAAGACGAGCTGAGACGACACCCGAGATGGCACTTCGATGCGCGACCGCTCTG 540
Db 857 CTGCAAGACGAGCTGAGACGACACCCGAGATGGCACTTCGATGCGCGACCGCTCTG 916
Qy 541 AGCGCGGTGAAGGAGCTCCACGAGACGAGACGAGACGAGATGAGAGATGCGATCCGAG 600
Db 917 AGCGCGGTGAAGGAGCTCCACGAGACGAGACGAGACGAGATGAGAGATGCGATCCGAG 976
Qy 601 GCAGAAAGCGAAGTGGATCAGATATCTCCGCGCGCATGAAATGCGAGACGCGGCG 660
Db 977 GCAGAAAGCGAAGTGGATCAGATATCTCCGCGCGCATGAAATGCGAGACGCGGCG 1036
Qy 661 AAAGTCTTAGTCCCGCGCGCTCTCCGCTTACCGATGCGCAACGCGCAACTCTT 720
Db 1037 AAAGTCTTAGTCCCGCGCGCTCTCCGCTTACCGATGCGCAACGCGCAACTCTT 1096
Qy 721 GTGGCGCGCGCGCATCCGAGACGATCAGACGACGACGACCACTGCACTGAGACGAA 780
Db 1097 GTGGCGCGCGCGCATCCGAGACGATCAGACGACGACGACCACTGCACTGAGACGAA 1156
Qy 781 CAAGTCAATCAGACGACGAACTGTCCAAAGTTTGAACAG 822

Db 1157 CAACTGCATCAGCAGCACTGTCCAAACAGTTTACGAACG 1198
|||||
RESULT 11
AB017577 535 bp DNA linear INV 11-NOV-1999
LOCUS Drosophila yakuba gene for extramacrochaetae, partial cds.
DEFINITION AB017577.1 GI:6331662
ACCESSION AB017577.1 GI:6331662
VERSION extramacrochaetae; emc.
KEYWORDS Drosophila yakuba
SOURCE Drosophila yakuba
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS Takano-Shimizu,T.
TITLE Local recombination and mutation effects on molecular evolution in
Drosophila
JOURNAL Genetics 153 (3), 1285-1296 (1999)
MEDLINE 20014821
PUBMED 10545459
REFERENCE 2 (bases 1 to 535)
AUTHORS Takano,T.S.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1998) Toshiyuki S. Takano, National Institute of
Genetics, Department of Population Genetics, Yata 111, Mishima,
Shizuoka 411-8540, Japan (E-mail:totakano@lab.nig.ac.jp,
Tel:81-559-81-6781, Fax:81-559-81-6784)
location/Qualifiers
FEATURES
source 1..535
/organism="Drosophila yakuba"
/mol_type="genomic DNA"
/db_xref="taxon:7245"
1..535
/gene="emc"
/gene="emc"
/gene="emc"
/codon_start=1
/product="extramacrochaetae"
/protein_id="BAAB6621.1"
/db_xref="GI:6331662"
/translation="AVCTGASGMPALNAGRIORHPTHRGDSNAEMKMYLSKLKDL
VPEPKRKLKLEIIOHVIDYICDLTLETHEMENPDAALTAANGHDEED9
MEDADAEAEVDADYLAGRLNAEQPAKVSPPARLPLTRQPTNTLVAAHPPQOOO
QOOLQOQOOLQSOOOLS"
BASE COUNT 133 a 183 c 149 g 70 t
ORIGIN
Query Match 24.7%; Score 503; DB 3; Length 535;
Best Local Similarity 96.3%; Pred. No. 9.6e-78;
Matches 515; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy 274 GCCGTCTGCCAGACAGTGCTCTCCGATGCGGCTTTGAATGCGAGCGGCATCCAG 333
Db 1 GCCGTCTGCCAGACAGTGCTCTCCGATGCGGCTTTGAATGCGAGCGGCATCCAG 60
Qy 334 CGCCATCCACGACCGCGCGAGCGGGAGAAACCGCGATGAAGATATCTCTCAA 393
Db 61 CGCCACCCACGACCGCGCGAGCGGGAGAAACCGCGATGAAGATATCTCTCAA 120
Qy 394 CTGAAGAGCTCGTTCGTTGATGCGCAAGAAAGCAAGCACTCAAGCTGAGATATC 453
Db 121 CTGAAGAGCTCGTTCGTTGATGCGCAAGAAAGCAAGCACTCAAGCTGAGATATC 180
Qy 454 CAGCAGCTGATGATCTGACATCTGCGACTGCGACCGGATGAGAGCGACCCCGAATG 513
Db 181 CAGCAGCTGATGATCTGACATCTGCGACTGCGACCGGATGAGAGCGACCCCGAATG 240
Qy 514 GGCACCTTGATGCGGAGCGGCTCTGAGCGCGGTGAACGAGCTCCAGAGAGAGAC 573
Db 241 GGCACCTTGATGCGGAGCGGCTCTGAGCGCGGTGAACGAGCTCCAGAGAGAGAC 300

Qy 574 AGGCATGAGAGATGCGATGCGGAGGAGAAAGTGCATTCAGTATCTCGCC 633
Db 301 AGGCATGAGAGATGCGATGCGGAGGAGAAAGTGCATTCAGTATCTCGCC 360
Qy 634 CAGCGCTGATATCCGAGCAGCGCGGAAAGTTCTTAATGCCGCGCTCTCCGCTT 633
Db 361 CAGCGCTGATATCCGAGCAGCGCGGAAAGTTCTTAATGCCGCGCTCTCCGCTT 420
Qy 694 ACCGATGCCAAGCGCCAACTCTTGTGCGCGCCCGCATCCGAGCAGATCAGAG 753
Db 421 ACCGATGCCAAGCGCCAACTCTTGTGCGCGCCCGCATCCGAGCAGATCAGAG 480
Qy 754 CAGCAGCACTGCACTGACGACGACAACTGCAATCAACAGCACTGTCCA 808
Db 481 CAGCAGCACTGCACTGACGACGACAACTGCAATCAACAGCACTGTCCA 535
RESULT 12
AF174666 449 bp DNA linear INV 07-DEC-2000
LOCUS Drosophila melanogaster emc[S058701] P element insertion region.
DEFINITION AF174666
ACCESSION AF174666
VERSION AF174666.1 GI:6531697
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS 1 (bases 1 to 449)
TITLE Prokopenko,S.N., He,Y., Lu,Y. and Bellen,H.J.
Mutations affecting the development of the peripheral nervous
system in Drosophila: a molecular screen for novel proteins
JOURNAL Genetics 156 (4), 1691-1715 (2000)
MEDLINE 20556139
PUBMED 11102367
REFERENCE 2 (bases 1 to 449)
AUTHORS Prokopenko,S.N. and Bellen,H.J.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Molecular and Human Genetics/HIMI, Baylor
College of Medicine, T634, Mail stop BCM235, One Baylor Plaza,
Houston, TX 77030, USA
location/Qualifiers
FEATURES
source 1..449
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/map="61D1-2"
/note="flanking the 3'-end of the emc[S058701] P element;
macrochaetae (emc) gene
Sequence recovery method: 2.5 kb A7 genomic fragment was
isolated by plasmid rescue with EcoRI restriction enzyme;
fragment was directly sequenced at the proximal end
(relative to the site of P element insertion)"
BASE COUNT 147 a 122 c 103 g 76 t 1 others
ORIGIN
Query Match 17.5%; Score 355.8; DB 3; Length 449;
Best Local Similarity 99.2%; Pred. No. 4.6e-52;
Matches 368; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 8 GTCTGACCGTCCGACGAAAGATATGAATGAGAACTTAACGCAAAAACAGCTG 67
Db 1 GTCTGACCGTCCGACGAAAGATATGAATGAGAACTTAACGCAAAAACAGCTG 60
Qy 68 GATTAAAGTCTTGGTTTGAACATGACGAGTGAGAAACAAAGTGAAGAGGCGAC 127
Db 61 GATTAAAGTCTTGGTTTGAACATGACGAGTGAGAAACAAAGTGAAGAGGCGAC 120
Qy 128 AAAGCCGAATCAATCAATTGAAGCAAGTCAAGAGCAATCTTAACTGAACCTCGC 187

Db 121 AAAGCCGAATCAATCAATTGAATAAGCAGTCAGACATCTTAACTGAACCTCGC 180
188 AAGTAATCCGCTGACGCGCTGTAAACAACCGCAAGATTCTACTCAAGACGCTTAACAA 247
181 AAGTAATCCGCTGACGCGCTGTAAACAACCGCAAGATTCTACTCAAGACGCTTAACAA 240
Qy 248 TCATCTCCAGAAATGAATGCTCCGTGACGCGCTGTGACAGAGTGTCTCCGGAATGCGG 307
241 TCATCTCCAGAAATGAATGCTCCGTGACGCGCTGTGACAGAGTGTCTCCGGAATGCGG 300
Db 308 CTTTGAATGCGAGCGGCGCATCCAGCGCATCCACGACCGCGGAGCGAG-AGAAC 366
301 CTTTGAATGCGAGCGGCGCATCCAGCGCATCCACGACCGCGGAGCGAGATTAAC 360
Qy 367 GCCGAGATGAA 377
361 GCCGAAATGAA 371
Db
RESULT 13
LOCUS AF049135 477 bp mRNA linear VRT 06-MAR-1998
DEFINITION Gallus gallus Id2 mRNA, complete cds.
ACCESSION AF049135
VERSION AF049135.1 GI:2935460
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 477)
Spirin,K.S., Kenney,M.C. and Koefler,H.P.
Extreme conservation of id genes among vertebrates
Unpublished
2 (bases 1 to 477)
Spirin,K.S., Kenney,M.C. and Koefler,H.P.
Direct Submission
Submitted (17-FEB-1998) Ophthalmology Research, Cedars-Sinai
Medical Center, 8700 Beverly Blvd., 2025 Davis Research Bldg., Los
Angeles, CA 90048, USA
FEATURES
source
1. 477
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/dev_stage="9 day embryo"
19. 423
/note="helix-loop-helix inhibitor of DNA binding"
/codon_start=1
/product="Id2"
/protein_id="AAC05130.1"
/db_xref="GI:2935461"
/translation="MKAFSPVRSVRKNGLSERNLGISRSKTPVDWPSLLYNNNDGYS
KKEIVPSIPQNKVSKMELIQLHVIDIYLDIQLALDSHPISIVSLHHQRPQNAASRTP
LTLTNDISILSLQAAFPSEBLMASDSKALCG"
BASE COUNT 104 a 160 c 121 g 92 t
ORIGIN
Query Match 4.2%; Score 85.6; DB 5; Length 477;
Best Local Similarity 71.8%; Pred. No. 6.6e-05;
Matches 112; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Db 249 CGCGTGAAGTCCGACCCGCTCCATGTCAGCTCCA 284
RESULT 14
LOCUS AF068831 1239 bp mRNA linear VRT 26-AUG-1998
DEFINITION Gallus gallus transcriptional regulator Id2 mRNA, complete cds.
ACCESSION AF068831
VERSION AF068831.1 GI:3192886
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1239)
Martinsen,B.J. and Bronner-Frazer,M.
Neural crest specification regulated by the helix-loop-helix
repressor Id2
JOURNAL Science 281 (5379), 988-991 (1998)
MEDLINE 98369172
PUBMED 9703514
REFERENCE 2 (bases 1 to 1239)
Martinsen,B.J. and Bronner-Frazer,M.
Direct Submission
Submitted (28-MAY-1998) Division of Biology, California Institute
of Technology, Pasadena, CA 92697, USA
FEATURES
source
1. 1239
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
76. 480
/function="DNA binding inhibitor"
/note="helix-loop-helix"
/codon_start=1
/product="transcriptional regulator Id2"
/protein_id="AAC32827.1"
/db_xref="GI:3192887"
/translation="MKAFSPVRSVRKNGLSERNLGISRSKTPVDWPSLLYNNNDGYS
KKEIVPSIPQNKVSKMELIQLHVIDIYLDIQLALDSHPISIVSLHHQRPQNAASRTP
LTLTNDISILSLQAAFPSEBLMASDSKALCG"
BASE COUNT 343 a 304 c 274 g 318 t
ORIGIN
Query Match 4.2%; Score 85.6; DB 5; Length 1239;
Best Local Similarity 71.8%; Pred. No. 7.3e-05;
Matches 112; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

REFERENCE

1 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Garcia, T., Roman Roman, S., Baron, R., Call, K., Theilhaber, J.,

TITLE

Genes involved in osteogenesis, and methods of use

JOURNAL

Patent: WO 02081745-A 61 17-OCT-2002;
Aventis Pharma S.A. (FR)

FEATURES

Location/Qualifiers
1..420

/organism="Mus musculus"
/mol_type="genomic DNA"

/db_xref="taxon:10090"

BASE COUNT

107 a 133 c 102 g 78 t

ORIGIN

Query Match 4.1%; Score 83.4; DB 6; Length 420;

Best Local Similarity 70.7%; Pred. No. 0.00016;

Matches 11; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATGTATCTGTCTCAACTGAAGACCTGTTCCGTTCAATGCCCAAGAACAG 428
Db 119 CAACATGACGACCTGCTACTCTCAAGCTCAAGAGAGCTGTGCTCCAGCATCCCCAGAACAA 178
QY 429 GAAGCTCACCAAGCTGAGATCATTCAGACGTCATGACTTACATCTGGCACTGCAAGAC 488
Db 179 GAAGTGAGCAAGATGGAATCCCTGCAAGACGTCATGACTTACATCTGGCACTGCAAGAT 238
QY 489 CGAGCTGAGAGCGACCCCGAGATGGGCACTTGAT 525
Db 239 CGCCTGACTGCATCCCACTATGTCTAGCTGCAAT 275

Search completed: February 8, 2004, 04:40:58
Job time : 7763.72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:33:25 ; Search time 580.083 Seconds
(without alignments)
9474.605 Million cell updates/sec

Title: US-09-614-150A-11
Perfect score: 2036
Sequence: 1 atcgtgtcgtcgcgcgtcg.....ttaataatgcacaaattt 2036

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N.GeneSeq.19Jun03:*

1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2001.DAT:*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT:*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2003.DAT:*
25: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2036	100.0	2036	23	ABLO1847
2	1215	59.7	6080	23	ABLO1846
3	83.4	4.1	420	25	ABZ34703
4	83.4	4.1	420	25	ABZ34704
5	83.4	4.1	530	24	ABK54695
6	83.4	4.1	532	24	ABK54758
7	83.4	4.1	533	24	ABK5068
8	83.4	4.1	1024	25	ABZ83203

9	83.4	4.1	1049	24	ABX04171	Human mRNA differe
10	83.4	4.1	1049	24	ABK83598	Human cDNA differe
11	83.4	4.1	1276	24	AB199305	Mouse ischaemic co
12	83.4	4.1	1348	21	AAFL6172	Human prostate can
13	83.4	4.1	2511	22	AAK89597	Genomic sequence #
14	83.4	4.1	2511	22	AAK89596	Human digestive sy
15	82.6	4.1	624	24	ABT09421	Phase-1 Rat CT gen
16	81.8	4.0	405	24	AAAD40553	Human Id2 DNA. Ho
17	81.8	4.0	640	24	ABU68617	Kidney cancer rela
18	81.8	4.0	640	25	ABZ34779	Coding sequence SE
19	80	3.9	481	24	ABZ35731	Human ID1 polynuci
20	80	3.9	481	24	ABX09974	Human ID1 DNA frag
21	80	3.9	481	24	ABV78155	Human ID1 DNA SRO
22	80	3.9	481	24	ABU91696	Human polynucleoti
23	80	3.9	483	25	ABZ34702	Coding sequence SE
24	80	3.9	509	15	AAO66082	Human Id-1H gene.
25	80	3.9	926	24	AAK16579	DNA encoding human
26	80	3.9	926	25	ABZ34777	Coding sequence SE
27	80	3.9	979	22	AAK44963	cDNA encoding nove
28	79.8	3.9	1216	20	AAZ33520	Human prostate can
29	79.8	3.9	4793	24	AAK16586	DNA encoding human
30	78.8	3.9	721	20	AAK40105	Gastric cancer ass
31	78.2	3.8	502	15	AAO66083	Human Id-1H' gene.
32	78.2	3.8	1167	24	ABN95937	Gene #2435 used to
33	78.2	3.8	1167	24	ABL68359	Kidney cancer rela
34	76.8	3.8	682	25	ABZ83034	Toxicologically re
35	73.8	3.6	2029	23	ABL07061	Drosophila melanog
36	73.8	3.6	6539	23	ABL07060	Drosophila melanog
37	72	3.5	365	24	ABK54848	Human colon cancer
38	72	3.5	927	24	AAK19635	Mouse ischaemic co
39	69.6	3.4	1027	21	AAK18245	Lung cancer associ
40	67.8	3.3	8056	25	ABZ10246	Haematopoietic cel
41	67.6	3.3	8056	25	ABZ10100	Haematopoietic cel
42	66.4	3.3	8056	25	ABZ10246	Haematopoietic cel
43	65.2	3.2	6109	24	ABJ33730	Human immune syste
44	65	3.2	486	24	ABZ35733	Human ID4 polynuci
45	65	3.2	486	24	ABX09976	Human ID4 DNA frag

ALIGNMENTS

RESULT 1	ABLO1847	standard; cDNA; 2036 BP.
ID	ABLO1847	
AC	ABLO1847;	
XX		
DT	26-MAR-2002 (first entry)	
XX		
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 23.	
XX		
KW	Drosophila; developmental biology; cell signalling; insecticide;	
KM	pharmaceutical; gene; ss.	
XX		
OS	Drosophila melanogaster.	
XX		
PN	WO200171042-A2.	
XX		
PD	27-SEP-2001.	
XX		
PF	23-MAR-2001; 2001WO-US09231.	
XX		
PR	23-MAR-2000; 2000US-191637P.	
XX		
PR	11-JUL-2000; 2000US-0614150.	
XX		
PA	(PEKE) PE CORP NY.	
XX		
PI	Venter JC, Adams M, Li PWD, Myers EW;	
XX		
DR	WPI; 2001-656860/75.	
XX		
DR	P-PSDB; ABB57744.	
XX		

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT interactions -

XX Claim 1; SEQ ID NO 23; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB5737-AB572072), and the encoded proteins

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published_pct_sequences.

SQ Sequence 2036 BP; 703 A; 478 C; 385 G; 470 T; 0 other;

Query Match 100.0%; Score 2036; DB 23; Length 2036;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 2036; Conservative 0; Indels 0; Gaps 0;

```
QY 1 ATTGTTGTCGACCGCTCGAGCAAAAGATATGAATCGAAGCGTAAACGAGAAA 60
DB 1 ATTGTTGTCGACCGCTCGAGCAAAAGATATGAATCGAAGCGTAAACGAGAAA 60
QY 61 CACGCTGATTAAGTGTTCGTTTCGATCATGACAGAGTGAGAAAACAAAGTAAAGT 120
DB 61 CACGCTGATTAAGTGTTCGTTTCGATCATGACAGAGTGAGAAAACAAAGTAAAGT 120
QY 121 AGGCGCAAAAGCCGAATCAAAATGAATGAAGCAAGCAAGTCTTAACTGA 180
DB 121 AGGCGCAAAAGCCGAATCAAAATGAATGAAGCAAGCAAGTCTTAACTGA 180
QY 181 ACCTCGCAAGTATCCGGTACGACCGTGTAAACAACCGCAAGATTCTTCAAGAGAC 240
DB 181 ACCTCGCAAGTATCCGGTACGACCGTGTAAACAACCGCAAGATTCTTCAAGAGAC 240
QY 241 AAAGCAATCATCTCCAGATGAAGTCCCTGACGGCCGTCTGCGACAGAGTGCCTCCGGA 300
DB 241 AAAGCAATCATCTCCAGATGAAGTCCCTGACGGCCGTCTGCGACAGAGTGCCTCCGGA 300
QY 301 ATGCGGCGCTTGAATGCGAGCGGGCGATCCAGGCGCATCCAGCAACCGGCGGAGCGG 360
DB 301 ATGCGGCGCTTGAATGCGAGCGGGCGATCCAGGCGCATCCAGCAACCGGCGGAGCGG 360
QY 361 GAGAACCCGAGATGAAGATGTATCTGTCCAACTGAAGGACCTGTTCGTTTCATGCC 420
DB 361 GAGAACCCGAGATGAAGATGTATCTGTCCAACTGAAGGACCTGTTCGTTTCATGCC 420
QY 421 AAGAACAGAAAGCTCACCAGCTGAGATCATCCAGACGTCATCGACTATCTGGGAC 480
DB 421 AAGAACAGAAAGCTCACCAGCTGAGATCATCCAGACGTCATCGACTATCTGGGAC 480
QY 481 CTGAGAACCGAGCTGAGAGCGCACCCGAGATGGGCAACTTCGATGGGCGACCGCTCTG 540
DB 481 CTGAGAACCGAGCTGAGAGCGCACCCGAGATGGGCAACTTCGATGGGCGACCGCTCTG 540
QY 541 ACAGGCGGTGAACGAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 ACAGGCGGTGAACGAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 GCAGAAACGGAAGTCGATTCAGATATCTCCGACGAGCGCTGAATGCGAGCAACCGGCG 660
DB 601 GCAGAAACGGAAGTCGATTCAGATATCTCCGACGAGCGCTGAATGCGAGCAACCGGCG 660
QY 661 AAATCTCTAGTCCCGCGCGCGCGCTCCCGTTACGATGGCGCAAAAGCCCAACTCTT 720
DB 661 AAATCTCTAGTCCCGCGCGCGCGCTCCCGTTACGATGGCGCAAAAGCCCAACTCTT 720
QY 721 GTGCGCGCGCGCGCGCGCGCGCTCCCGTTACGATGGCGCAAAAGCCCAACTCTT 780
DB 721 GTGCGCGCGCGCGCGCGCGCGCTCCCGTTACGATGGCGCAAAAGCCCAACTCTT 780
```

```
DB 721 GTGCGCGCGCGCGCGCGCGCGCTCCCGTTACGATGGCGCAAAAGCCCAACTCTT 780
QY 781 CAACTGCAATCAGAGAGCAACTGTCCAAAGTTTGAACAGCGCAGAGATGGGAGAAA 840
DB 781 CAACTGCAATCAGAGAGCAACTGTCCAAAGTTTGAACAGCGCAGAGATGGGAGAAA 840
QY 841 GACAGAGACAGTGTGAAGGTGGAACCGCCAGCGGAGAAAACCAAGCAAGCTATTAAG 900
DB 841 GACAGAGACAGTGTGAAGGTGGAACCGCCAGCGGAGAAAACCAAGCAAGCTATTAAG 900
QY 901 CTTATTAATTAATGATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 960
DB 901 CTTATTAATTAATGATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 960
QY 961 AGATCTCCGATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1020
DB 961 AGATCTCCGATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1020
QY 1021 TTCAAGCGCTCTTTTCCCGGAGAAACCGGAGAAAACCGAGACTTCGCTTTGGCT 1080
DB 1021 TTCAAGCGCTCTTTTCCCGGAGAAACCGGAGAAAACCGAGACTTCGCTTTGGCT 1080
QY 1081 CCAACAGAGTTTGTATTAACATTTAATAACAAGATTAACAACAACAACAACAACA 1140
DB 1081 CCAACAGAGTTTGTATTAACATTTAATAACAAGATTAACAACAACAACAACAACA 1140
QY 1141 AGAGCGCGGTAAAGTACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
DB 1141 AGAGCGCGGTAAAGTACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
QY 1201 GCGCTTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1260
DB 1201 GCGCTTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1260
QY 1261 TTTAAGCTTGAATTAACATGAAGCAATATTTTGTTCCTAATTTTAACTTTAGTT 1320
DB 1261 TTTAAGCTTGAATTAACATGAAGCAATATTTTGTTCCTAATTTTAACTTTAGTT 1320
QY 1321 TGGTTTAAACAAGCAAGAAAGAAAGTAACTAATGAAGCAAGAGATGAGCAATAA 1380
DB 1321 TGGTTTAAACAAGCAAGAAAGAAAGTAACTAATGAAGCAAGAGATGAGCAATAA 1380
QY 1381 AAAAAAATTAATAAATGACAGAGCAACAACCTGAATTAAGTTGATCAACGTTTCTG 1440
DB 1381 AAAAAAATTAATAAATGACAGAGCAACAACCTGAATTAAGTTGATCAACGTTTCTG 1440
QY 1441 CATTAGATTAATGAATTTCAATCCAAATGATTTTCTATTAATTTTAACTGAAGAG 1500
DB 1441 CATTAGATTAATGAATTTCAATCCAAATGATTTTCTATTAATTTTAACTGAAGAG 1500
QY 1501 TTTTCCATTAATGATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1560
DB 1501 TTTTCCATTAATGATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1560
QY 1561 ACAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
DB 1561 ACAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
QY 1621 AAGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
DB 1621 AAGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
QY 1681 TCAATTTGATTTTCCGAGTCCCTGAGTCTTTCCTTTTAACTGAATTAAGCG 1740
DB 1681 TCAATTTGATTTTCCGAGTCCCTGAGTCTTTCCTTTTAACTGAATTAAGCG 1740
QY 1741 ATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
DB 1741 ATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
QY 1801 GTTCTGTGATCTGATGATTTTGAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1860
DB 1801 GTTCTGTGATCTGATGATTTTGAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1860
```

Db	1801	GTTCTGTCATCTGAGCATACCTCTTGACAAAGTGTGTTTTTTTAGACATAAATTATGCAATTT	1860
Qy	1861	TAATTATATCATCCGCAATGGAATTTACACATTTTGTCTCATTTTCGATGGTAAAAAACACT	1920
Db	1861	TAATTATATCATCCGCAATGGAATTTACACATTTTGTCTCATTTTCGATGGTAAAAAACACT	1920
Qy	1921	CTATCATGATGTAATAAAAAAAAAAACAAGAAGAAACATTTGATGATGTTGCCAAAAAATAT	1980
Db	1921	CTATCATGATGTAATAAAAAAAAAAACAAGAAGAAACATTTGATGATGTTGCCAAAAAATAT	1980
Qy	1981	ATTATCTTAAACATTAAGAAAAACAACACTAGCTCTTAATTAATATTCGAAAAATTTT	2036
Db	1981	ATTATCTTAAACATTAAGAAAAACAACACTAGCTCTTAATTAATATTCGAAAAATTTT	2036

RESULT 2

ID	AB101846 standard; cDNA; 6080 BP.
XX	AB101846;
AC	
XX	
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 20.
XX	
KM	Drosophila; developmental biology; cell signalling; insecticide;
XX	pharmaceutical; gene; ss.
OS	
XX	Drosophila melanogaster.
PN	
XX	MO200171042-A2.
PD	
XX	27-SEP-2001.
PF	
XX	23-MAR-2001; 2001WO-US09231.
XX	
XX	23-MAR-2000; 2000US-191637P.
PR	
XX	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	
XX	Venter JC, Adams M, Li FWD, Myers EW;
DR	
XX	WPI; 2001-6556860/75.
DR	
XX	P-PSDB; ABB57743.
PT	
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
XX	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	
XX	Claim 1; SEQ ID NO 20; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (AB116176-ABL30511), expressed DNA
CC	sequences (AB101840-ABL16175) and the encoded proteins
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp://who.int/pub/published_pct_sequences.
XX	
SQ	Sequence 6080 BP; 1785 A; 1353 C; 1216 G; 1726 T; 0 other;
Query Match	59.7%; Score 1215; DB 23; Length 6080;
Best Local Similarity	100.0%; Pred. No. 2.8e-239;
Matches 1215; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	822 GCCACAGATTCGGAGAAAGCAGCAGCACTCGTAAAGCTGGGAAAGCCCGCGGAA 881
DB	3866 GCCACAGATTCGGAGAAAGCAGCAGCACTCGTAAAGCTGGGAAAGCCCGCGGAA 3925

QY	882	AAACCAGGAAACCTTAATGAAGCTTAACTAAATGATATGATATGAAAGTACAGTCAAAACCTCC	941
Db	3926	AAACCAGGAAACCTTAATGAAGCTTAACTAAATGATATGAAAGTACAGTCAAAACCTCC	3998
QY	942	ACTCGCAACCTTCGTGTGAAGATCTCCGCATTAAGATTAAGTGTGAACCTCTCAAGAAAC	1001
Db	3986	ACTCGCAACCTTCGTGTGAAGATCTCCGCATTAAGATTAAGTGTGAACCTCTCAAGAAAC	4045
QY	1002	CACAGCCACGGCGGATCTCTTCAAGCGCTTTTCCGGGAAAGGAAACCCGGGAAA	1061
Db	4046	CACAGCCACGGCGGATCTCTTCAAGCGCTTTTCCGGGAAAGGAAACCCGGGAAA	4105
QY	1062	CGAGACCTCGCTTTTGGCTCCACAGAGTTAGTTTAAACATTTTAAACAAGATACAA	1121
Db	4106	CGAGACCTCGCTTTTGGCTCCACAGAGTTAGTTTAAACATTTTAAACAAGATACAA	4165
QY	1122	CAACCAAAAACCTTGCATTTAGAGCAGGTTAAGCAATGACATTAATGATTTTAAATTA	1181
Db	4166	CAACCAAAAACCTTGCATTTAGAGCAGGTTAAGCAATGATTAATTAATTAATTA	4225
QY	1182	CATTACAACTAGATGATGCGCTTTAAGATGAACGGAAGAAAGAAACCTTGCGAAAGCT	1241
Db	4226	CATTACAACTAGATGATGCGCTTTAAGATGAACGGAAGAAAGAAACCTTGCGAAAGCT	4285
QY	1242	TTAAACAATTTTAAATTTTCCCTTAAACGCTTGAATTAACACTGAAGCATTAATTTTGT	1301
Db	4286	TTAAACAATTTTAAATTTTCCCTTAAACGCTTGAATTAACACTGAAGCATTAATTTTGT	4345
QY	1302	CTTAATTTTGGTTTGGTTTGGTTTAAACAACGGAAGAAAGAAAGTTTAAATGCAAGC	1361
Db	4346	CTTAATTTTGGTTTGGTTTGGTTTAAACAACGGAAGAAAGTTTAAATGCAAGC	4405
QY	1362	AAACAGAGATCAGCAATTAACAAAAATTAATTAATGAATGACAGGACCAACCTGAAATTA	1421
Db	4406	AAACAGAGATCAGCAATTAACAAAAATTAATTAATGAATGACAGGACCAACCTGAAATTA	4465
QY	1422	GTTGAATCAACGTTTCTCGCATTAACGATTAATGATTTCAATCCAAATGATTTTCTAAT	1481
Db	4466	GTTGAATCAACGTTTCTCGCATTAACGATTAATGATTTCAATCCAAATGATTTTCTAAT	4525
QY	1482	AAATTTATCATCAATGAAGTTTTCCTTAATTAATGATTAATGACACGACCAATTAACAACAGA	1541
Db	4526	AAATTTATCATCAATGAAGTTTTCCTTAATTAATGATTAATGACACGACCAATTAACAACAGA	4585
QY	1542	ACAACCAACTCAGCAGACACACACATTCAAATTCATTAATTTAATTCAGAGAGTACGT	1601
Db	4586	ACAACCAACTCAGCAGACACACACATTCAAATTCATTAATTTAATTCAGAGAGTACGT	4645
QY	1602	TTTAAGTTTAAGTTTCGAGAAAGTTCAATTTCCAGGCACTAATTTTCACTCAACACACTAC	1661
Db	4646	TTTAAGTTTAAGTTTCGAGAAAGTTCAATTTCCAGGCACTAATTTTCACTCAACACACTAC	4705
QY	1662	ATAGAAAAACAACAACACTCAATTTTGTAGTTTCCGAGTCCCTTGATCGTTTTCCTCT	1721
Db	4706	ATAGAAAAACAACAACACTCAATTTTGTAGTTTCCGAGTCCCTTGATCGTTTTCCTCT	4765
QY	1722	TTTTTAATCGAATAACGCAATATGCAATTAATTAATTAATTCACAGATTAACCTAAACCTA	1781
Db	4766	TTTTTAATCGAATAACGCAATATGCAATTAATTAATTAATTCACAGATTAACCTAAACCTA	4825
QY	1782	TAAATTAATTAATTAATTAATGTTTCGATATCTGTGACATCTTTGTACAAAGTGTTTT	1841
Db	4826	TAAATTAATTAATTAATTAATGTTTCGATATCTGTGACATCTTTGTACAAAGTGTTTT	4885
QY	1842	TAGCATTAATTAATGCAATTTTAATTAATGATCCGCAATCGAAATTTACATTTTGTCTCA	1901
Db	4886	TAGCATTAATTAATGCAATTTTAATTAATGATCCGCAATCGAAATTTACATTTTGTCTCA	4945
QY	1902	TTTCGATGTTAAACAACCTCATATCATATGTTAAATTAATTAATTAATTAATTAATTTGA	1961
Db	4946	TTTCGATGTTAAACAACCTCATATCATATGTTAAATTAATTAATTAATTAATTAATTTGA	5005

QY 1962 TGAATGTCGCAAAATATATATATCTAAACATTAAGAAAAACAACACTAGCTTTAAT 2021
DB 5006 TGAATGTCGCAAAATATATATATCTAAACATTAAGAAAAACAACACTAGCTTTAAT 5065
QY 2022 AAATTGCAAAATTT 2036
DB 5066 AAATTGCAAAATTT 5080

RESULT 3

ABZ34703
ID ABZ34703 standard; cDNA; 420 BP.

AC ABZ34703;

DT 04-FEB-2003 (first entry)

DE Coding sequence SEQ ID 61, upregulated in osteogenesis.

XX Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;

KM osteoporosis; bone disease; downregulator; human;

KW transcription factor; orphan nuclear receptor; ss.

XX Homo sapiens.

XX MO200281745-A2.

PD 17-OCT-2002.

PF 05-APR-2002; 2002WO-IB02211.

PR 05-APR-2001; 2001US-281400P.

PA (AVET) AVENTIS PHARMA SA.

PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;

PI Connolly T, Jackson A, Bushnell SE, Rawadi G;

DR WPI; 2003-058567/05.

PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,

PT useful for bone disease therapy in subject

XX Claim 26; Page 109; 237pp; English.

XX The present invention relates to novel nucleotide sequences, which are

CC differentially expressed in models of osteogenesis upon being put in

CC contact with a stimulator of osteogenesis. The present sequence is one

CC such sequence. This sequence can be used for diagnosing osteoporosis/bone

CC disease in a patient, promoting osteogenesis and/or preventing

CC osteoporosis/bone disease. The present sequence encodes a transcription

CC factor or an orphan nuclear receptor.

XX Sequence 420 BP; 107 A; 133 C; 102 G; 78 T; 0 other;

QY Query Match 4.1%; Score 83.4; DB 25; Length 420;

DB Best Local Similarity 70.7%; Pred. No. 1.4e-07;

Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAGATGATCTGTCGCAAACTGAGAGACTCGTTCCGTTCAATGCCAAGAACG 428

DB 119 CAACATGAGAGACTGCTACTCTCAAGCTCAAGAGCTGGTCCGAGCATCCCGAGAACAA 178

QY 429 GAAGCTCACCAGCTGAGATCATCCAGACGTCATGATCATCTGCCAGCTCCGACAG 488

DB 179 GAAGGTGAGCAAGATGGAATCTCGACGACGTCATGATCATCTTGACCTGACGAT 238

ABZ34704
ID ABZ34704 standard; cDNA; 420 BP.
XX
AC ABZ34704;

DT 04-FEB-2003 (first entry)

DE Coding sequence SEQ ID 62, upregulated in osteogenesis.

XX Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;

KM osteoporosis; bone disease; downregulator; human;

KW transcription factor; orphan nuclear receptor; ss.

XX Homo sapiens.

XX MO200281745-A2.

PD 17-OCT-2002.

PF 05-APR-2002; 2002WO-IB02211.

PR 05-APR-2001; 2001US-281400P.

PA (AVET) AVENTIS PHARMA SA.

PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;

PI Connolly T, Jackson A, Bushnell SE, Rawadi G;

DR WPI; 2003-058567/05.

PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,

PT useful for bone disease therapy in subject

XX Claim 26; Page 109; 237pp; English.

XX The present invention relates to novel nucleotide sequences, which are

CC differentially expressed in models of osteogenesis upon being put in

CC contact with a stimulator of osteogenesis. The present sequence is one

CC such sequence. This sequence can be used for diagnosing osteoporosis/bone

CC disease in a patient, promoting osteogenesis and/or preventing

CC osteoporosis/bone disease. The present sequence encodes a transcription

CC factor or an orphan nuclear receptor.

XX Sequence 420 BP; 107 A; 133 C; 102 G; 78 T; 0 other;

QY Query Match 4.1%; Score 83.4; DB 25; Length 420;

DB Best Local Similarity 70.7%; Pred. No. 1.4e-07;

Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAGATGATCTGTCGCAAACTGAGAGACTCGTTCCGTTCAATGCCAAGAACG 428

DB 119 CAACATGAGAGACTGCTACTCTCAAGCTCAAGAGCTGGTCCGAGCATCCCGAGAACAA 178

QY 429 GAAGCTCACCAGCTGAGATCATCCAGACGTCATGATCATCTGCCAGCTCCGACAG 488

DB 179 GAAGGTGAGCAAGATGGAATCTCGACGACGTCATGATCATCTTGACCTGACGAT 238

QY 489 CGAGCTGAGAGAGCAACCCCGAGATGGGCAACTTCGAT 525

DB 239 CGCCTGAGCTGCAATCCCACTATTGTTCAGCCTGCAT 275

RESULT 5

ABK54695
ID ABK54695 standard; cDNA; 530 BP.

AC ABK54695;

DT 18-JUN-2002 (first entry)

DE Human colon cancer-associated cDNA, SEQ ID No 165.

XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX

XX Homo sapiens.
OS
XX
XX WO200212280-A2.
PN
XX
PD 14-FEB-2002.
XX
XX 30-JUL-2001; 2001WO-US23826.
XX
XX 03-AUG-2000; 2000US-223265P.
PR 02-OCT-2000; 2000US-237406P.
PR 20-MAR-2001; 2001US-277495P.
PR 03-JUL-2001; 2001US-302702P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Pyle RA, Xu J, Secrist H;
PI
XX WPI; 2002-257462/30.
XX
XX Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers
XX
XX Claim 1; Page 198; 425pp; English.
XX
XX The invention relates to isolated polynucleotides (i) encoding colon
CC tumour polypeptides (ii). (i) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (i) are useful for determining the presence
CC of cancer in a patient. (i) and (ii) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (ii) is useful for inhibiting
CC development of cancer in a patient. (i) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (i). ABK54531-ABK55464 represent human colon cancer cDNA
CC sequences of the invention.
XX
SQ Sequence 530 BP; 124 A; 172 C; 138 G; 96 T; 0 other;

Query Match 4.1%; Score 83.4; DB 24; Length 530;
Best Local Similarity 70.7%; Pred. No. 1.5e-07;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATGTATCTGTCCAACTGAAGACCTGTTCCGTTATGCCCAAGAACAG 428
DB 218 CAACATGAAGCACTGCTACTCCAGCTCAAGCTCAAGAGCTGTGCTCCAGCATCCCGAAGAACAA 277
QY 429 GAAGCTCACCAAGCTGAGATCATCCAGCAGTCATGACATCTGCGACCTGCGAGAC 488
DB 278 GAAGGTGAGCAAGATGAAATCTCGACGACGTCATGACATCTTGAGACCTGCGAGAT 337
QY 489 CGAGCTGAGAGCGACCCCGAGATGGGCAACTTCGAT 525
DB 338 CGCCCTGAGACTCGCATCCCACTATTGTCAAGCTGCGAT 374

RESULT 6
ABK54758
ID ABK54758 standard; cDNA; 532 BP.
XX
XX ABK54758;
AC
XX 18-JUN-2002 (first entry)
XX
XX Human colon cancer-associated cDNA, SEQ ID No 228.
DE
XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
OS
XX Homo sapiens.
XX
XX WO200212280-A2.
PN

XX 14-FEB-2002.
PD
XX
XX 30-JUL-2001; 2001WO-US23826.
XX
XX
XX 03-AUG-2000; 2000US-223265P.
PR 02-OCT-2000; 2000US-237406P.
PR 20-MAR-2001; 2001US-277495P.
PR 03-JUL-2001; 2001US-302702P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Pyle RA, Xu J, Secrist H;
PI
XX WPI; 2002-257462/30.
XX
XX Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers
XX
XX Claim 1; Page 215; 425pp; English.
XX
XX The invention relates to isolated polynucleotides (i) encoding colon
CC tumour polypeptides (ii). (i) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (i) are useful for determining the presence
CC of cancer in a patient. (i) and (ii) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (ii) is useful for inhibiting
CC development of cancer in a patient. (i) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (i). ABK54531-ABK55464 represent human colon cancer cDNA
CC sequences of the invention.
XX
SQ Sequence 532 BP; 124 A; 174 C; 139 G; 95 T; 0 other;

Query Match 4.1%; Score 83.4; DB 24; Length 532;
Best Local Similarity 70.7%; Pred. No. 1.5e-07;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATGTATCTGTCCAACTGAAGACCTGTTCCGTTATGCCCAAGAACAG 428
DB 221 CAACATGAAGCACTGCTACTCCAGCTCAAGAGCTGTGCTCCAGCATCCCGAAGAACAA 280
QY 429 GAAGCTCACCAAGCTGAGATCATCCAGCAGTCATGACATCTTGCGACCTGCGAGAC 488
DB 281 GAAGGTGAGCAAGATGAAATCTCGACGACGTCATGACATCTTGAGACCTGCGAGAT 340
QY 489 CGAGCTGAGAGCGACCCCGAGATGGGCAACTTCGAT 525
DB 341 CGCCCTGAGACTCGCATCCCACTATTGTCAAGCTGCGAT 377

RESULT 7
ABK55068
ID ABK55068 standard; cDNA; 533 BP.
XX
XX ABK55068;
AC
XX 18-JUN-2002 (first entry)
XX
XX Human colon cancer-associated cDNA, SEQ ID No 538.
DE
XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.
OS
XX Homo sapiens.
XX
XX WO200212280-A2.
PN
XX 14-FEB-2002.
XX
XX 30-JUL-2001; 2001WO-US23826.
PF

XX 03-AUG-2000; 2000US-223265P.
PR 02-OCT-2000; 2000US-237406P.
PR 20-MAR-2001; 2001US-277495P.
PR 03-JUL-2001; 2001US-302702P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Pyle RA, Xu J, Secretist H;
XX
XX WPI; 2002-257462/30.
XX
XX Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers -
XX
XX
PS Claim 1; Page 304; 425pp; English.
XX
XX The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (I) are useful for determining the presence
CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (II) is useful for inhibiting
CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). ABK54531-ABK55464 represent human colon cancer CDNA
CC sequences of the invention.
XX
SQ Sequence 533 BP; 124 A; 173 C; 138 G; 94 T; 4 other;
Query Match 4.1%; Score 83.4; DB 24; Length 533;
Best Local Similarity 70.7%; Pred. No. 1.5e-07;
Matches 11; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 369 CGAGATGAGATGATCTGTCCAACTGAGAGACCTGTCCTTCATGCCCAAGACAG 428
DB 221 CAACATGAAGACTGCTACTCTCAAGCTCAGAGAGCTGCTCCAGCATCCCCAGAAACAA 280
QY 429 GAAGCTCACCAGCTGGAGATCATCCAGCAGTCATGCTGCACTTGGCAGCTGCAGAC 488
DB 281 GAAGGTGAGCAGATGAGAAATCTGACAGCAGTCATGATGATCTTGGACCTGCAGAT 340
QY 489 CGAGCTGAGAGCAGACCCCGAGATGGGCACTTGAT 525
DB 341 CGCCTGAGCTCGCATCCCACTATTGTTCAGCTGCAT 377
RESULT 8
ABZ83203
ID ABZ83203 standard; cDNA; 1024 BP.
XX
XX ABZ83203;
XX
XX 14-MAY-2003 (first entry)
XX
XX Toxicologically relevant human nucleotide sequence #362.
XX
XX Toxicologically relevant gene; toxicological response; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2003016500-A2.
XX
XX 27-FEB-2003.
XX
XX 16-AUG-2002; 2002WO-US26514.
XX
XX 16-AUG-2001; 2001US-313080P.
XX
XX (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.

XX Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schweiser K;
PI Allen P;
XX
XX WPI; 2003-268322/26.
XX
XX Determining a toxicological response to an agent, useful for screening
PT of drugs, comprises comparing the expression profile of one or more
PT human toxic response genes to a reference gene expression profile
PT indicative of toxicity -
XX
XX
PS Claim 1; Page 132; 455pp; English.
XX
XX The present invention describes a method (M1) for determining a
CC toxicological response to an agent, which comprises comparing the
CC expression profile of one or more human toxic response genes to a
CC reference gene expression profile indicative of toxicity, and so
CC determining the presence of a toxic response to the agent. Also
CC described: (1) an array comprising one or more polynucleotides selected
CC from the genes corresponding to the partial sequences given in ABZ82842
CC to ABZ84764, or their fragments of at least 20 nucleotides, or
CC homologues; and (2) determining if a gene putatively identified to be a
CC toxic response gene plays a role on toxic response pathways by
CC determining the expression profile of the gene after exposure of cells
CC or a human subject to a known toxic pharmaceutical or industrial agent,
CC comprising: (a) exposing cells to an agent or isolating cells from a
CC human subject who was exposed to an agent; (b) obtaining the test gene
CC expression profile for a putatively identified toxic response gene after
CC exposure to a known toxic pharmaceutical or industrial agent; and
CC (c) comparing the test profile to the expression profile of a gene with
CC a similar function or comparing the test profile to the expression
CC profile of that gene after exposure to other known toxic compounds. The
CC methods are useful for predicting and determining toxicological responses
CC on a cellular, organ or system level. The arrays comprising the human
CC genes are useful for toxicological screening of drugs, pharmaceutical
CC compounds and chemicals.
XX
SQ Sequence 1024 BP; 290 A; 254 C; 217 G; 263 T; 0 other;
Query Match 4.1%; Score 83.4; DB 25; Length 1024;
Best Local Similarity 70.7%; Pred. No. 1.8e-07;
Matches 11; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 369 CGAGATGAGATGATCTGTCCAACTGAGAGACCTGTCCTTCATGCCCAAGACAG 428
DB 207 CAACATGAAGACTGCTACTCTCAAGCTCAGAGAGCTGCTCCAGCATCCCCAGAAACAA 266
QY 429 GAAGCTCACCAGCTGGAGATCATCCAGCAGTCATGCTGCACTTGGCAGCTGCAGAC 488
DB 267 GAAGGTGAGCAGATGAGAAATCTGACAGCAGTCATGATGATCTTGGACCTGCAGAT 326
QY 489 CGAGCTGAGAGCAGACCCCGAGATGGGCACTTGAT 525
DB 327 CGCCTGAGCTCGCATCCCACTATTGTTCAGCTGCAT 363
RESULT 9
ABX04171
ID ABX04171 standard; cDNA; 1049 BP.
XX
XX ABX04171;
XX
XX 10-JAN-2003 (first entry)
XX
XX Human mRNA differentially expressed in mesenchymal cells #18.
XX
XX Human; ss; gene; skeletal growth; cartilage degeneration disorder;
XX chondroblastic phenotype; mesenchymal cell; cartilage formation;
XX bone formation; arthritis; osteoarthritis; rheumatoid arthritis;
XX gout arthritis; adjuvant arthritis; arthritis deformans; osteoporosis;
XX infectious arthritis; osteochondritis; RDA; antiarthritic; osteopathic;
XX antirheumatic; antiinflammatory; representational difference analysis.

OS Homo sapiens.
 XX
 XX WO200271927-A2.
 XX
 XX
 PD 19-SEP-2002.
 XX
 XX 12-MAR-2002; 2002WO-US07787.
 XX
 XX 12-MAR-2001; 2001US-274980P.
 XX
 XX (BGM) BRIGHAM & WOMENS HOSPITAL INC.
 PA Yates KE, Mizuno S, Glowacki J;
 PI WPI; 2002-723276/78.
 XX
 XX
 DR New nucleic acid molecules capable of promoting chondrogenesis, useful
 PT for diagnosing and treating cartilaginous tissue degeneration
 PT conditions, e.g., osteoarthritis, rheumatoid arthritis, gout arthritis,
 PT or osteochondrosis -
 XX
 XX Claim 33; Page 114; 153pp; English.
 PS
 XX The invention relates to new isolated nucleic acid molecule comprising a
 CC nucleic acid molecule consisting of a gene differentially expressed
 CC in cells undergoing differentiation from mesenchymal cell to a
 CC chondroblastic phenotype, or hybridising under stringent conditions
 CC to them (or their fragments). Also included are expression vectors,
 CC transformed host cells, expressed polypeptides or peptide fragments
 CC (which induce differentiation of a mesenchymal cell and may be used as an
 CC immunogen), binding partners of the polypeptides, a method for
 CC identifying an agent useful in modulating mesenchymal cell
 CC differentiation induction activity of a molecule, a method of diagnosing
 CC a condition characterized by aberrant expression of a nucleic acid
 CC molecule or its expression product; a method for determining regression,
 CC progression or onset of cartilaginous tissue degeneration condition in a
 CC subject characterized by aberrant expression of a nucleic acid molecule
 CC or its expression product; a method for treating a cartilaginous tissue
 CC degeneration condition; a method for treating a subject to reduce the
 CC risk of cartilaginous tissue degeneration condition developing in the
 CC subject; a method for identifying a candidate agent for treating a
 CC cartilaginous tissue degeneration condition; and a solid-phase nucleic
 CC acid molecule array consisting essentially of a set of nucleic acid
 CC molecule as cited above (or known from known genes shown to be
 CC differentially expressed in developing mesenchymal cells using the
 CC technique of representational difference analysis, RDA), its expression
 CC products or fragments, fixed to a solid substrate. The nucleic acids,
 CC polypeptides and agents are useful for treating cartilaginous tissue
 CC degeneration conditions such as osteoarthritis, rheumatoid arthritis,
 CC gout arthritis, adjuvant arthritis, arthritis deformans, infections
 CC arthritis or osteochondrosis. The present invention is a cDNA from
 CC a known gene differentially expressed in developing mesenchymal cells.
 XX
 XX Sequence 1049 BP; 298 A; 258 C; 222 G; 271 T; 0 other;
 SQ
 Query Match 4.1%; Score 83.4; DB 24; Length 1049;
 Best Local Similarity 70.7%; Pred. No. 1.9e-07;
 Matches 11; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 369 CGAGATGAGATGATGTCCTCAACAGGAGAGCTGCTCCGTTCCGCCAAGAACAG 428
 DB 207 CAACATGAGAGACTGCTACTCCAGCTCAAGAGCTGCTCCGATCCGCCAAGAAC 266
 QY 429 GAAGCTCACCAAGCTGAGATCATTCGACAGCTGATGACTGACACTGCGACAG 488
 DB 267 GAAGGTGAGCAAGATGGAATCTCGACAGCACTGATGATCATCTTGAGACCTGCGAGAT 326
 QY 489 CGAGCTGAG 525
 DB 327 CGCCCTGAGCTCGAGCTCCACTATGTCAGCCCTGCGAT 363

ABK83598
 ID ABK83598 standard; cDNA; 1049 BP.
 XX
 XX AC ABK83598;
 XX
 XX 14-AUG-2002 (first entry)
 XX
 XX Human cDNA differentially expressed in granulocytic cells #169.
 DE
 XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 XX Homo sapiens.
 OS
 XX WO200228999-A2.
 XX
 XX 11-APR-2002.
 PD
 XX 03-OCT-2001; 2001WO-US30821.
 XX
 XX 03-OCT-2000; 2000US-237189P.
 PR
 XX (GENE-) GENE LOGIC INC.
 PA Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 PI WPI; 2002-435328/46.
 DR
 XX Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -
 XX
 XX Claim 1; SEQ ID No 169; 114pp; English.
 PS
 XX The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating GA; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease, also bacterial infection, viral infection,
 CC parasitic infection, protozoal infection, fungal infection and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1049 BP; 298 A; 258 C; 222 G; 271 T; 0 other;

Query Match 4.1%; Score 83.4; DB 24; Length 1049;
Best Local Similarity 70.7%; Pred. No. 1.9e-07;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATGATCTGTCCTCAAACTGAAGAGACCTGTTCCGTTATGCCAAGAACG 428
DB 207 CAACATGACGACCTGCTACTCCAGCTCAAGAGACCTGGTCCGACATCCCCGAGAACAA 266
QY 429 GAAGCTGACCAAGCTGGAGATCATCCAGACCTGATGATCATCTGCGACCTGCAGAC 488
DB 267 GAAAGTGACCAAGATGAAATCTGCGACGACGTCATGACATCTTGGACCTGCAGAT 326
QY 489 CGAGCTGAGAGACGACCCCGAGATGGGCAACTTCGAT 525
DB 327 CGCCTGAGACTGCGATCCCATATTGTCAGCCTGCAT 363

RESULT 11
ABI99305
ID ABI99305 standard; cDNA; 1276 BP.

AC ABI99305;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:160.

KM Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

OS Mus musculus.

PN W0200188188-A2.

PD 22-NOV-2001.

PF 18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

PI Iehikawa K, Asai S, Takahashi Y, Nagata T, Iehi Y;

DR WPI; 2002-034733/04.

PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -

PS Claim 2; Page 472; 2690pp; English.

XX The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,

CC which are used in the exemplification of the present invention.

SQ Sequence 1276 BP; 372 A; 282 C; 257 G; 365 T; 0 other;

Query Match 4.1%; Score 83.4; DB 24; Length 1276;
Best Local Similarity 70.7%; Pred. No. 2e-07;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATGATCTGTCCTCAAACTGAAGAGACCTGTTCCGTTATGCCAAGAACG 428
DB 187 CAACATGACGACCTGCTACTCCAGCTCAAGAGACCTGGTCCGACATCCCCGAGAACAA 246
QY 429 GAAGCTGACCAAGCTGGAGATCATCCAGACCTGATGATCATCTGCGACCTGCAGAC 488
DB 247 GAAAGTGACCAAGATGAAATCTGCGACGACGTCATGACATCTTGGACCTGCAGAT 306
QY 489 CGAGCTGAGAGACGACCCCGAGATGGGCAACTTCGAT 525
DB 307 CGCCTGAGACTGCGATCCCATCTTGTGACGCTGCAT 343

RESULT 12
AAF16172
ID AAF16172 standard; cDNA; 1348 BP.

AC AAF16172;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:607.

KM Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

KW neuroprotective; cytostatic; cardioprotective; immunomodulatory; muscular;

KW vulnerary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;

KW antibacterial; gene therapy; neural; immune; reproductive; renal;

KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KW wound; infectious disease; ss.

OS Homo sapiens.

PN W020005174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI (ROSE/) ROSEN C A.

DR Rosen CA, Ruben SM;

DR WPI; 2000-587513/55.

DR P-PSDB; AAB56969.

PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -

PS Claim 1; Page 1055; 2338pp; English.

XX AAF1566 to AAF1650s encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioprotective, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX

Sequence 1348 BP; 402 A; 306 C; 271 G; 368 T; 1 other;

Query Match 4.1%; Score 83.4; DB 21; Length 1348;
Best Local Similarity 70.7%; Pred. No. 2e-07;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATATATCTGTCCAACTGAGAGACCTGCTTCACTGACCAAGACAG 428
DB 237 CACATGAAACGACTGCTACTCTCAAGCTCAAGAGAGCTGTGCTCCAGCATCCCGAACA 296
QY 429 GAAGCTCACCAGCTGAGATCATTCACAGACGCTCATGATGATGACCTGAGAC 488
DB 297 GAAGGTGAGCAATGAAATCTCGACAGACGTCATGACTATGATCTGAGACCTGAGAT 356
QY 489 CGAGCTGAGAGCCACCCCGAGATGGCACTTCGAT 525
DB 357 CGCCCTGAGACTGCATCCCATTTGTCAGCTCGAT 393

RESULT 13

AA539597/c
ID AA539597 standard; DNA; 2511 BP.

XX AA539597;

XX 17-DEC-2001 (first entry)

DE Genomic sequence #16 encoding human colon associated polypeptide.

XX Human; colon cancer; congenital abnormality; infection; colitis;
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;
KW sigmoid disease; antibacterial; antiviral; antiinflammatory;
XX cytostatic; ds.

XX Homo sapiens.

PN WO200155302-A2.

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01240.

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249287.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251859.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
DR WPI; 2001-465567/50.
XX
PT Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the colon including colon cancers and also for
PT testing and detection e.g. diagnosis -
XX
PS Disclosure; SEQ ID No 494; 562bp; English.
XX
XX The present invention relates to the isolation of novel human colon
XX associated polypeptides (AAU22468-AAU22701), and the cDNA and genomic
XX sequences encoding for them. The sequences of the invention are useful
XX in the diagnosis, treatment, prevention and/or prognosis of disorders
XX of the colon including colon cancer, congenital abnormalities
XX (e.g. atresia and stenosis), bacterial and viral infections,
XX inflammatory bowel disease (IBD), neoplastic cell disorders,
XX (e.g. polyps and adenomas), intestinal inflammatory disorders, colitis,
XX colonic inflammation, diarrhoea and dysentery, malabsorption syndromes
XX (e.g. lactose intolerance), intestinal obstruction and sigmoid diseases.
XX The polynucleotide sequences of the invention can also be used in gene
XX therapy. AA839582-AA84060 represent DNA sequences encoding for the
XX novel human colon associated polypeptides of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 2511 BP; 730 A; 548 C; 586 G; 647 T; 0 other;

Query Match 4.1%; Score 83.4; DB 22; Length 2511;
Best Local Similarity 70.7%; Pred. No. 2.5e-07;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 369 CGAGATGAAGATGATATCTGTCCAACTGAGAGACCTGTTCCGTTGATGCCAAGACAG 428
Db 2179 GACATGACGACTGCTACTCCAACTCAAGAGCTGGGCCAGATCCCAAGAACAA 2120
Qy 429 GAAGCTCACCAGCTGAGATCATTCAGACCGTCATCATCTTGTGGAACCTGAGAC 488
Db 2119 GAAGGTGAGCAGAAATGAAATCTGACAGACGTCATCATCTTGTGGAACCTGAGAT 2060
Qy 489 CGAGCTGAGAGACGACCCGAGATGGGCAACTTCGAT 525
Db 2059 CGCCTGAGCTCCATCCCATCTTGTCAAGCTGAT 2023
RESULT 14
AAK8996/c
ID AAK8996 standard; DNA; 2511 BP.
XX
XX AAK8996;
AC
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 2572.
XX
XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX digestive system disorder; Meckel's diverticulum; ds.
XX
XX Homo sapiens.
XX
XX WO20015514-A2.
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01324.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 14-AUG-2000; 2000US-0226279.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226682.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.

Patent No. 5527897
GENERAL INFORMATION:
APPLICANT: Oda, Kinichiro
APPLICANT: Nakada, Susumu
APPLICANT: Hara, Eiji
APPLICANT: Yamaguchi, Tomoko
APPLICANT: Nakamura, Takeshi
APPLICANT: Oka, Yumiko
APPLICANT: Kishimoto, Toshiniko
TITLE OF INVENTION: Human ID Genes
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
STREET: 99 Canal Center Plaza, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/151,391A
FILING DATE: 12-NOV-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mills, Demetra J.
REGISTRATION NUMBER: 34,506
REFERENCE/DOCKET NUMBER: 715-087
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)684-1111
TELEFAX: (703)684-1124
TELEX: 82-4425
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 502 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc feature
LOCATION: 22..465
OTHER INFORMATION: /note= "CDS"
US-08-151-391A-3

Query Match Similarity	3.9%;	Score 79.8;	DB 1;	Length 502;
Best Local Similarity	65.4%;	Pred. No. 3.3e-08;		
Matches 117;	Conservative 0;	Mismatches 62;	Indels 0;	Gaps 0.

QY	345	GCACCGCGCGCA	CGGGGAGAA	CGCCGAGATGA	TGATGTTC	CCAACTGAAGACCT	404
Db	195	GCAGAGGTAA	CGTGTCTCTA	GCACATGA	CGCGTACTAC	CGCTCAAGAGCT	254
QY	405	CGTTCGTTCA	TGCCAAGAAC	GAGAGTCA	CCCACTGGAGAT	CAATCCAGCACTCAT	464
Db	255	GGTGCCCA	CCCTGCCCA	AACTCGCAAG	TGAGGAGAT	TCTCCACACATCAT	314
QY	465	CGACTCATCT	TCGCAC	CTGCAGAC	CGAGCTGGAG	ACGCCCGAGATGGGCAATTTGG	523
Db	315	CGACTCATCA	AGGAGCTT	TCAGTTGAG	CTGAACTCGGAAT	CCGAATTGGAAATCCCCG	373

RESULT 3
US-09-561-497-10
; Sequence 10, Application US/09561497
; Patent No. 637243
; GENERAL INFORMATION:
; APPLICANT: Brenda F. Baker

```

? APPLICANT: C. Frank Bennett
? APPLICANT: Jacqueline Wyatt
? TITLE OF INVENTION: ANTISENSE MODULATION OF INHIBITOR OF DNA BINDING-1 EXPRESSION
? FILE REFERENCE: RTS-0149
? CURRENT APPLICATION NUMBER: US/09/561,497
? CURRENT FILING DATE: 2000-04-28
? NUMBER OF SEQ ID NOS: 88
? SEQ ID NO 10
? LENGTH: 4793
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (2210)...(2659)
US-09-561-497-10

```

	Query Match	Best Local Match	Similarity	Score	79.8 ; DB 4 ;	Length	4793 ;
	Matches	117 ;	Conservative	0 ;	Mismatches	62 ;	Indels
							Gaps
							0 ;
Qy	345	GCAACCGCGGACCGGAGNAAACGCCGAGATGAAGATGTATCTGTCCAAACTGAAGACCT	404				
Db	2386	GCACGAGGTAAACGTGTGCTCTACGACATGAAACGGCTGTATCTACGACCTCAAGAGACT	2445				
Qy	405	CGTTCCGTTCAATGCCCAAGAACGAAAGCTCAACCACTGGAGATCATCCAGCAACCTCAT	464				
Db	2446	GGTCCCAACCTCGCCCCAGAACCCCAAGAGTGAGCAAGTGGAGATTCTCCAGCAGCTCAT	2505				
Qy	465	CGACTATCATCTGCGACCTCGAGACCGAGCTGGAGACGCAACCCCGAGATGGGCAACTTCG	523				
Db	2506	CGACTATCATCGGAGACCTTTCAGTTGAGAGCTGAATCTCGAATCCGAAATGGAAACCCCG	2564				

RESULT 4
 US-08-896-164-30
 Sequence 30, Application US/08896164
 Patent No. 6218521
 GENERAL INFORMATION:
 APPLICANT: OBARA, Yuichi
 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
 WITH GASTRIC CANCER AND METHODS FOR
 TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
 NUMBER OF SEQUENCES: 87
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Felfe & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 COUNTRY: USA
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/896,164
 FILING DATE: July 17, 1997
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6218521man D. Hanson
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-3884
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 721 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-896-164-30

Query Match 3.9%; Score 78.8; DB 3; Length 721;
Best Local Similarity 64.8%; Pred. No. 6.1e-08;
Matches 116; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 345 GCACCGCGGCGACGGGAGAACCCGAGATGATGTATCTCTCAAACTGAAGACCT 404
DB 239 GCAGCAGGTAAACGTCTCTCTCAACATGAACGGCTTACTCAAGCTCAAGAGCT 298
QY 405 CGTTCGCTTCATGCCCCAAGAACGAACTCACCAGCTGGAGATCATCCAGACGTCAT 464
DB 299 GGTGCCACCTGCCCCAGAACCGCAAGGTGAGCAGATGTAATCTCCAGACGTCAT 358
QY 465 CGACTACATCTGCGACCTGCGACGACCGAGCTGAGAGCACCAGATGGCAACTTGG 523
DB 359 CGACTACATCAGGAGACCTTCAGTTGAGCTGAACTCGGATCCGAGTTGGAACCCCG 417

RESULT 5
US-08-151-391A-1

; Sequence 1, Application US/08151391A

; Patent No. 5527897

; GENERAL INFORMATION:

; APPLICANT: Oda, Kinichiro

; APPLICANT: Nakada, Susumu

; APPLICANT: Hara, Eiji

; APPLICANT: Yamaguchi, Tomoko

; APPLICANT: Nakamura, Takeshi

; APPLICANT: Oka, Yumiko

; APPLICANT: Kishimoto, Toshihiko

; TITLE OF INVENTION: Human ID Genes

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER

; STREET: 99 Canal Center Plaza, Suite 300

; CITY: Alexandria

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22314

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/151,391A

; FILING DATE: 12-NOV-1993

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Miller, Demetra J.

; REGISTRATION NUMBER: 34,506

; REFERENCE/DOCKET NUMBER: 715-087

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)684-1111

; TELEFAX: (703)684-1124

; TELEX: 82-4425

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 509 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; FEATURE: misc.feature

; NAME/KEY: 22.483

; LOCATION: 22.483

; OTHER INFORMATION: /note= "CDS"

Query Match 3.9%; Score 78.4; DB 1; Length 509;
Best Local Similarity 58.1%; Pred. No. 6.5e-08;
Matches 158; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 345 GCACCGCGGCGACGGGAGAACCCGAGATGATGTATCTCTCAAACTGAAGACCT 404
DB 195 GCAGCAGGTAAACGTCTCTCTCAACATGAACGGCTTACTCAAGCTCAAGAGCT 254
QY 405 CGTTCGCTTCATGCCCCAAGAACGAACTCACCAGCTGGAGATCATCCAGACGTCAT 464
DB 255 GGTGCCACCTGCCCCAGAACCGCAAGGTGAGCAGATGTAATCTCCAGACGTCAT 314
QY 465 CGACTACATCTGCGACCTGCGACGACCGAGCTGAGAGCACCAGATGGCAACTTGA 524
DB 315 CGACTACATCAGGAGACCTTCAGTTGAGCTGAACTCGGATCCGAGTTGGAACCCCG 374
QY 525 TGGGCGAG--CGCTCTGACGGCGGTGAACGGACTTCACAGAGACGAGACGCACT 581
DB 375 GGGCCGAGGGCTGCGCGCTCCGGCTCTCCGCTCAGCACCTTCAACGGGAGATCAGCCCT 434
QY 582 GGAGGATGGGATGCCGAGGAGCAAGCGGAG 613
DB 435 GACGCCGAGGGCGGCGATGCGCTCCCTGCGAGC 466

RESULT 6
US-09-086-663A-78

; Sequence 78, Application US/0908663A

; Patent No. 6518063

; GENERAL INFORMATION:

; APPLICANT: DUCY, PATRICIA

; APPLICANT: KARSENTY, GERARD

; TITLE OF INVENTION: OSF2/CFP1 COMPOSITIONS AND METHODS OF USE

; FILE REFERENCE: UTSC:525

; CURRENT APPLICATION NUMBER: US/09/086,663A

; CURRENT FILING DATE: 1998-05-29

; PRIOR APPLICATION NUMBER: 60/080,189

; PRIOR FILING DATE: 1998-03-24

; PRIOR APPLICATION NUMBER: 60/048,430

; PRIOR FILING DATE: 1997-05-29

; NUMBER OF SEQ ID NOS: 83

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 78

; LENGTH: 2156

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (498)..(2060)

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; US-09-086-663A-78

Query Match 2.9%; Score 58.2; DB 4; Length 2156;
Best Local Similarity 64.4%; Pred. No. 0.0023;
Matches 87; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 666 CTCTAGTCCGCGCGCTCTCCGCTTACCATGCGCAAGCCCAACTCTTGTGC 725
DB 575 CTTCAGCCCCCTCCAGCAGCCTGCGAGCCCGCAAAATGAGAGCGTGCCTGTG 634
QY 726 GCGCGCCCATCGGCGACATCAGAGCAGAGCAACTGCAATGCAAGCAACT 785
DB 635 GGTGCGCAACGAGCAGCAACAGAGCAGAGCAAGCAAGCAGCAGCAGCAACA 694
QY 786 GCAATCAGACAGA 800
DB 695 GCAAGCAGCAGCA 709

RESULT 7
US-08-997-685A-1

; Sequence 1, Application US/08997685A

```

: Patent No.6551821
:
: GENERAL INFORMATION:
:
: APPLICANT: The Trustees of Columbia University
:
: APPLICANT: Kandell, Eric
:
: TITLE OF INVENTION: Brain Cyclic Nucleotide Gated Ion Channel and Uses Thereof
:
: FILE REFERENCE: 0575/54806
:
: CURRENT APPLICATION NUMBER: US/08/997,685A
:
: CURRENT FILING DATE: 1997-12-12
:
: NUMBER OF SEQ ID NOS: 60
:
: SOFTWARE: Patenlin version 3.1
:
: SEQ ID NO 1
:
: LENGTH: 2733
:
: TYPE: DNA
:
: ORGANISM: mouse
:
: US-08-997-685A-1

```

Query Match	2.7%	Score 55	DB 4	Length 2733
Best Local Similarity	66.4%	Pred. No.	0.012	
Matches 79	Conservative 0	Mismatches 40	Indels 0	Gaps 0

[illegible]

Qy	793	CAGCAGCAACGTCTCAACAGTTT	TAGCAGACCCACAGATTCGAGAAAGACACGAGAGA	851
Db	2320	CAGCAGCCACAGACACCTGTGTACTCACA	CCGAAAATGAACTGCACACAGACACACA	2378

RESULT 8
US-09-620-312D-475

```

Sequence 475, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Weinman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillingshast
APPLICANT: Dmanac, Radjic T.
TITLE OF INVENTION: Polyelectronic Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784C1P2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/468,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pf_fl_genes Version 1.0
SEQ ID NO 475
LENGTH: 3302
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (98)..(2563)
US-09-620-312D-475

```

Query Match	2.6%;	Score 53.2;	DB 4;	Length 3302;
Best Local Similarity	82.4%;	Pred. No. 0.031;		

	Matches	61; Conservative	0; Mismatches	13; Indels	0; Gaps	0;
QY	731	CCCATCCGCGAGCAGCATTCAGCAGCAGCAATCTGCAT		CACGAGCAACAATGCAT		790
Db	450	CTCACCGACGACGACGACGACGACGACGACGACGAA		CAGCAACGACGACGACGACGAC		509
QY	791	CACGACGCAACTG				804
Db	510	AGCAGCGACGACG				523

RESULT 9
US-09-135-994-1
; Sequence 1, Application US/09135994A

```

!! Patent NO. 6280938
!! GENERAL INFORMATION:
!! APPLICANT: Rannum et al.
!! TITLE OF INVENTION: SCAT GENE AND METHODS OF USE
!! FILE REFERENCE: University of Minnesota
!! CURRENT APPLICATION NUMBER: US/09/135,994A
!! CURRENT FILING DATE: 1998-08-18
!! EARLIER APPLICATION NUMBER: 60/056,170
!! EARLIER FILING DATE: 1997-08-19
!! NUMBER OF SEQ ID NOS: 14
!! SOFTWARE: PatentIn Ver. 2.0
!! SEQ ID NO 1
!!
!! LENGTH: 477
!! TYPE: DNA
!! ORGANISM: Homo sapiens
!! US-09-135-994-1

```

Query Match	2.6%	Score 52.2	DB 3	Length 477
Best Local Similarly	77.8%	Pred. No.	0.026	
Matches 63	Conservative	0	Mismatches 18	Indels 0
				Gaps 0

Dy 724 GCGGCGGCCCATCCGCAGCAGCATACGACGACGCAATCTGCACCTGCAGACGACA 783

Ddb 115 GCGGCGGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 174

QY	784	CTGCAATCAGCAGCAACTG	804
Db	175	CAGCAGCAGCAGCAGCAG	195

RESULT 10
US-09-684-843A-1
; Sequence 1, Application US/09684843A

```

: Patent No. 6514755
: GENERAL INFORMATION:
: APPLICANT : Rannu et al.
: TITLE OF INVENTION: SCAT GENE AND METHODS OF USE
: FILE REFERENCE: Regents of the university of Minnesota
: CURRENT APPLICATION NUMBER: US/09/684,843A
: CURRENT FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/056,170
: PRIOR FILING DATE: 1997-08-19
: PRIOR APPLICATION NUMBER: 09/335,994
: PRIOR FILING DATE: 1998-08-18
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 477
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-684-843A-1

```

Query Match	2.6%	Score 52.2	DB 4	Length 477
Best Local Similarity	77.8%	Pred. No. 0.026		
Matches 63	Conservative	0	Mismatches 18	Indels 0
				Gaps 0

QY 724 GCGCGCCCATTCGAGCAGCAGCATTCGACGAGCAAA 783
| | | | | | | | | | | | | | | | | | | | | |
Db 115 GCGCGCCCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 174

Qy 784 CTGCAATCAGCAGCAACTG 804
Db 175 CAGCAGCAGCAGCAGCAG 195

RESULT 11

US-09-491-356C-7
; Sequence 7, Application US/09491356C
; Patent No. 6566061
; GENERAL INFORMATION:
; APPLICANT: Philibert, Robert A.
; APPLICANT: Gims, Edward I.
; APPLICANT: Delisle, Lynn
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13
; FILE REFERENCE: 9465.6US11
; CURRENT APPLICATION NUMBER: US/09/491.356C
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: PCT/US99/09365
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 60/083,465
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 6558
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-491-356C-7

Query Match

Best Local Similarity 2.6%; Score 52.2; DB 4; Length 6558;
Matches 81; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 676 GCCGCCCCGTCTCCGCTTACCGATCGCCAAACGCCCAACACTTGTGGCGCCGCCCAT 735
Db 6055 GCCCAGGGTGTCCAGCAGCAGCGTCCGCTTCAACTTCCCTGCTGAGCAGCAACAA 6114
Qy 736 CCCCAGCAGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 795
Db 6115 CAAACAGCAGCAACAAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 6174
Qy 796 CAGCAACTG 804
Db 6175 CAGCAGCAG 6183

RESULT 12

US-08-256-077-3
; Sequence 3, Application US/08256077
; Patent No. 5654188
; GENERAL INFORMATION:
; APPLICANT: Elimeier, Wilfried
; APPLICANT: Weith, Andreas
; TITLE OF INVENTION: Neuroblastoma-Associated Regulator Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; City: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,077
; FILING DATE: 23-JUNE-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.1360000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-256-077-3

Query Match

Best Local Similarity 2.5%; Score 51.6; DB 1; Length 360;
Matches 81; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 367 GCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 426
Db 124 GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183
Qy 427 AGGAGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 486
Db 184 ACTGAGCTTACGAGCTGAGGAAATCTTACAGCGCGGTGATGATGATGATGATGATGAT 243
Qy 487 ACCGAGCTG 496
Db 244 GTAGTCTTG 253

RESULT 13

US-08-466-127-3
; Sequence 3, Application US/08466127
; Patent No. 5683878
; GENERAL INFORMATION:
; APPLICANT: Elimeier, Wilfried
; APPLICANT: Weith, Andreas
; TITLE OF INVENTION: Neuroblastoma-Associated Regulator Gene
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; City: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,127
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.1360002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS

LOCATION: 1..357
US-08-466-127-3

Query Match 2.5%; Score 51.6; DB 1; Length 360;
Best Local Similarity 62.3%; Pred. No. 0.032;
Matches 81; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 367 GCCGAGTGAAGATGTATCTGTCCAACTGAAGACCTGTCGTTCCATGCCCAAGAAC 426
DB 124 GAGCAGTGAACCACTGCTACTCCCGCTGGGGAATCGTACCCGAGATCCGAGAGGC 183
QY 427 AGGAGCTCACCAAGCTGAGATCATCCAGACGTCATGACTATCGGACCTGCAG 486
DB 184 ACTCAGCTTAGCCAGGCTGGAATCTCTACAGCGGCTCATCGACTATTCGACCTGCAG 243
QY 487 ACCGAGCTGC 496
DB 244 GTAGTCCTGG 253

RESULT 14

US-08-256-077-1
; Sequence 1, Application US/08256077
; Patent No. 5654188
; GENERAL INFORMATION:
; APPLICANT: Elmeler, Wilfried
; TITLE OF INVENTION: Neuroblastoma-Associated Regulator Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,077
; FILING DATE: 23-JUNE-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.1360000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 982 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-256-077-1

Query Match 2.5%; Score 51.6; DB 1; Length 982;
Best Local Similarity 62.3%; Pred. No. 0.046;
Matches 81; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 367 GCCGAGTGAAGATGTATCTGTCCAACTGAAGACCTGTCGTTCCATGCCCAAGAAC 426
DB 247 GAGCAGTGAACCACTGCTACTCCCGCTGGGGAATCGTACCCGAGATCCGAGAGGC 306
QY 427 AGGAGCTCACCAAGCTGAGATCATCCAGACGTCATGACTATCGGACCTGCAG 486
DB 307 ACTCAGCTTAGCCAGGCTGGAATCTCTACAGCGGCTCATCGACTATTCGACCTGCAG 366

QY 487 ACCGAGCTGC 496
DB 367 GTAGTCCTGG 376

RESULT 15

US-08-466-127-1
; Sequence 1, Application US/08466127
; Patent No. 5683878
; GENERAL INFORMATION:
; APPLICANT: Elmeler, Wilfried
; TITLE OF INVENTION: Neuroblastoma-Associated Regulator Gene
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,127
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.1360002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 982 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 37..481
; US-08-466-127-1

Query Match 2.5%; Score 51.6; DB 1; Length 982;
Best Local Similarity 62.3%; Pred. No. 0.046;
Matches 81; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 367 GCCGAGTGAAGATGTATCTGTCCAACTGAAGACCTGTCGTTCCATGCCCAAGAAC 426
DB 247 GAGCAGTGAACCACTGCTACTCCCGCTGGGGAATCGTACCCGAGATCCGAGAGGC 306
QY 427 AGGAGCTCACCAAGCTGAGATCATCCAGACGTCATGACTATCGGACCTGCAG 486
DB 307 ACTCAGCTTAGCCAGGCTGGAATCTCTACAGCGGCTCATCGACTATTCGACCTGCAG 366
QY 487 ACCGAGCTGC 496
DB 367 GTAGTCCTGG 376

Search completed: February 8, 2004, 10:31:24
Job time : 129.784 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2004, 04:41:38 ; Search time 766.996 Seconds
(without alignments)
9778.248 Million cell updates/sec

Title: US-09-614-150A-11

Perfect score: 2036
Sequence: 1 atcgtctgcgtacgcgtcg.....ttaataatgcaaaattt 2036

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCRUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1977	97.1	2031	14	US-10-108-605-206 Sequence 206, App
2	1940.4	95.3	2241	14	US-10-108-605-116 Sequence 116, App
3	83.4	4.1	481	11	US-09-918-995-1165 Sequence 1165, A
4	83.4	4.1	530	10	US-09-919-580-165 Sequence 165, App
5	83.4	4.1	532	10	US-09-919-580-228 Sequence 228, App
6	83.4	4.1	533	10	US-09-919-580-538 Sequence 538, App
7	83.4	4.1	1049	13	US-10-096-534-30 Sequence 30, App1
8	83.4	4.1	1348	10	US-09-925-300-607 Sequence 607, App1
9	83.4	4.1	1397	13	US-10-115-831-69 Sequence 69, App1
10	83.4	4.1	2511	11	US-09-764-872-494 Sequence 494, App1
11	81.8	4.0	405	15	US-10-025-170-1 Sequence 1, App1
12	80	3.9	979	12	US-10-291-172-44 Sequence 44, App1
13	80	3.9	1162	13	US-10-119-428-47 Sequence 47, App1
14	79.8	3.9	663	13	US-10-027-632-141641 Sequence 141641, App1
15	79.8	3.9	663	14	US-10-027-632-141641 Sequence 141641, App1

16	79.8	3.9	958	12	US-10-453-351-1 Sequence 1, App1
17	78.8	3.9	721	9	US-09-835-992A-30 Sequence 30, App1
18	78.2	3.8	1167	10	US-09-880-107-2434 Sequence 2434, App1
19	72	3.5	365	10	US-09-919-580-318 Sequence 318, App1
20	69.6	3.4	1027	9	US-09-925-302-264 Sequence 264, App1
21	66.6	3.3	3673778	13	US-10-312-841-2 Sequence 2, App1
22	65.2	3.2	6109	13	US-10-311-455-1703 Sequence 1703, App1
23	65	3.2	616	13	US-10-029-386-22934 Sequence 22934, App1
24	65	3.2	3378	15	US-10-102-806-304 Sequence 304, App1
25	65	3.2	3907	12	US-10-264-048-205 Sequence 205, App1
26	64.4	3.2	1124	10	US-09-917-800A-1574 Sequence 1574, App1
27	64.4	3.2	1124	12	US-10-388-934-206 Sequence 206, App1
28	63.4	3.1	756	13	US-09-814-353-6226 Sequence 6226, App1
29	63.2	3.1	756	13	US-09-814-353-12504 Sequence 12504, App1
30	63.2	3.1	319	9	US-09-925-299-405 Sequence 405, App1
31	63.2	3.1	319	11	US-09-925-299-405 Sequence 405, App1
32	63.2	3.1	8467	13	US-10-311-455-81 Sequence 81, App1
33	62.4	3.1	14023	13	US-10-311-455-2078 Sequence 2078, App1
34	62	3.0	8170	13	US-10-240-453-131 Sequence 131, App1
35	61.8	3.0	11805	13	US-10-311-455-1722 Sequence 1722, App1
36	61.8	3.0	3673778	13	US-10-312-841-1 Sequence 1, App1
37	61.4	3.0	17594	13	US-10-311-455-1899 Sequence 1899, App1
38	61	3.0	5690	13	US-10-240-453-80 Sequence 80, App1
39	61	3.0	5690	15	US-10-239-676-74 Sequence 74, App1
40	60.2	3.0	6591	13	US-10-311-455-1224 Sequence 1224, App1
41	60.2	3.0	11577	13	US-10-311-455-413 Sequence 413, App1
42	59.8	2.9	2000	10	US-09-938-842A-5370 Sequence 5370, App1
43	59.8	2.9	2000	12	US-09-938-842A-5370 Sequence 5370, App1
44	59.4	2.9	6161	13	US-10-311-455-384 Sequence 384, App1
45	59.4	2.9	15387	13	US-10-311-455-157 Sequence 157, App1

ALIGNMENTS

RESULT 1
US-10-108-605-206
; Sequence 206, Application US/10108605
; Publication No. US20020160934A1
GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Steam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 206
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-10-108-605-206

Query Match 97.1%; Score 1977; DB 14; Length 2031;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2020; Conservative 0; Mismatches 10; Indels 7; Gaps 3;

QY 1 ATTCGTTGTCGACCGCTCGCAGCAAAAGATATGAAATCGAAGCGAGAA 60
DB 1 ATTCGTTGTCGACCGCTCGCAGCAAAAGATATGAAATCGAAGCGAGAA 60
QY 61 CACGCTGATTAAGTGTTCGTTTGAACATCGACAGGTAGAAACAAAGTGAAGT 120
DB 61 CACGCTGATTAAGTGTTCGTTTGAACATCGACAGGTAGAAACAAAGTGAAGT 120

Oy	1201	GGCGTTAAGATGAACGAGAGAAAGAAATCTTGCGAAAGCTTTAAACAATTAATTTCC	1266
Oy	1202	AGCGCAAAACCCGAAATCAAAATCAAAATTTGAAAAGCCAGTCGAAGGCAATCTTTAACTGA	180
Db	121	AGCGCAAAACCCGAAATCAAAATCAAAATTTGAAAAGCCAGTCGAAGGCAATCTTTAACTGA	180
Oy	181	ACCTCGCAAGTAATCCGGTAGACCGCTGTAAACAACCGCAAGATTTCTACTCAAGAGCGT	240
Db	181	ACCTCGCAAGTAATCCGGTAGACCGCTGTAAACAACCGCAAGATTTCTACTCAAGAGCGT	240
Oy	241	AAAGCAATCATCTCCAGATGAAGTCCCTGACGCGCGTCTGCAAGACAGGTGCTCCGGA	300
Db	241	AAAGCAATCATCTCCAGATGAAGTCCCTGACGCGCGTCTGCAAGACAGGTGCTCCGGA	300
Oy	301	ATGCGCGCTTGAANTGCCAGCGGGGCGCATCCAGGCGCATCCCAAGCACCCGGGAGAGCGG	360
Db	301	ATGCGCGCTTGAANTGCCAGCGGGGCGCATCCAGGCGCATCCCAAGCACCCGGGAGAGCGG	360
Oy	361	GAGAAACGCGAGATGATGATGTATCTGTCAAACTGAAAGACCTTGTTCCCTTCATGCCC	420
Db	361	GAGAAATGCCAGATGATGATGTATCTGTCAAACTGAAAGACCTTGTTCCCTTCATGCCC	420
Oy	421	AAAGAACAGAAAGCTTCAACCACTGGAGATCATCCAGACGTCATCGACTCATCTGCGAC	480
Db	421	AAAGAACAGAAAGCTTCAACCACTGGAGATCATCCAGACGTCATCGACTCATCTGCGAC	480
Oy	481	CTGCAAGACCGAGCTGAGAACGCCACCCGAGATGGGCAACTTCGATGGCGGACCGCTCTG	540
Db	481	CTGCAAGACCGAGCTGAGAACGCCACCCGAGATGGGCAACTTCGATGGCGGACCGCTCTG	540
Oy	541	ACGCGCGGTGAACGGACTCCACGAGGACGAGACAGCGCATGAGAGATGCGGATGCGAG	600
Db	541	ACGCGCGGTGAACGGACTCCACGAGGACGAGACAGCGCATGAGAGATGCGGATGCGAG	600
Oy	601	GCAGAAAGGGAAGTGCATCCAGATATCTCTGCGCCAGCGCTGAAATGCCAGACCGGCG	660
Db	601	GCAGAAAGGGAAGTGCATCCAGATATCTCTGCGCCAGCGCTGAAATGCCAGACCGGCG	660
Oy	661	AAAGTCTTAAATCCCGCGCGCGCTCTCCCGCTTAACCGATGCCCAAAAGCCCAACACTTT	720
Db	661	AAAGTCTTAAATCCCGCGCGCGCTCTCCCGCTTAACCGATGCCCAAAAGCCCAACACTTT	720
Oy	721	GTGGCGCCCGCCATCCGACAGCATCAGCAGCAGCAGCACTGCAACTCAGACGAA	780
Db	721	GTGGCGCCCGCCATCCGACAGCATCAGCAGCAGCAGCACTGCAACTCAGACGAA	780
Oy	781	CAACTGCATCAACAGCAGCAACTGTCCAAAGTTTACCAACGCCACAGAAATGCCGAGAA	840
Db	781	CAACTGCATCAACAGCAGCAACTGTCCAAAGTTTACCAACGCCACAGAAATGCCGAGAA	840
Oy	841	GACAGCAGACAGTGCCTAAGCGTGGAAAGCCCAAGCCGAGAAACCCACAGAACTCTAATAG	900
Db	841	GACAGCAGACAGTGCCTAAGCGTGGAAAGCCCAAGCCGAGAAACCCACAGAACTCTAATAG	900
Oy	901	CCTATTAACTAAATCATATGAAGTAAGTCAAGTCAAACTCCACTGCGCAACTTGGTGA	960
Db	901	CCTATTAACTAAATCATATGAAGTAAGTCAAGTCAAACTCCACTGCGCAACTTGGTGA	960
Oy	961	AGATCTCCGCAATAAGATAACGTGTGAACCTGTCAAGAAACCAACAGCCAGCGCATCTC	1020
Db	961	AGATCTCCGCAATAAGATAACGTGTGAACCTGTCAAGAAACCAACAGCCAGCGCATCTC	1020
Oy	1021	TTCAACGCTCTCTTTCCGGGAAACCGGAAACAGAGACTCGCTTTGGCT	1080
Db	1021	TTCAACGCTCTCTTTCCGGGAAACCGGAAACAGAGACTCGCTTTGGCT	1080
Oy	1081	CCAACAGAGTTAGTTAATTAACAATTAATCAAGAAATCAACACAGAAACCTAAGACTT	1140
Db	1081	CCAACAGAGTTAGTTAATTAACAATTAATCAAGAAATCAACACAGAAACCTAAGACTT	1140
Oy	1141	AGAGCAGCGTAAAGCAAGTACACATTATGATTTAATTTAAACATTACAACTAGATGATG	1200
Db	1141	AGAGCAGCGTAAAGCAAGTACACATTATGATTTAATTTAAACATTACAACTAGATGATG	1200
Oy	1201	GGCGTTAAGATGAACGAGAGAAAGAAATCTTGCGAAAGCTTTAAACAATTAATTTCC	1266

Db	1201	GCCTTAAAGATGACGGAGAAAGAAACCTTGGCAAGCTTTAAACAATTAATTTTCC	1260
Qy	1261	TTTAACCTTGAAATTAACACTGAAGCATATATTTTGTTCCTAATTTAGTTTGT	1320
Db	1261	TTTAACCTTGAAATTAACACTGAAGCATATATTTTGTTCCTAATTTAGTTTGT	1320
Qy	1321	TGGTTTAAACAGCAGAAAGAAAATGTTACTAATGCAACAAACAAGATACAGCAAT -A	1379
Db	1321	TGGTTTAAACAGCAGAAAGAAAATGTTACTAATGCAACAAACAAGATACAGCAAT	1380
Qy	1380	AAAAAAATTAATAAAATGACAGGCAACCAACTGAAATTAAGTTAATCAGCTTCT	1439
Db	1381	AAAAAAATTAATAAAATGACAGGCAATCAACT ---TTACGTTGATCAGCGTTTCT	1436
Qy	1440	GCATTACGATTAATATGATTTCAATCCAAATGATTTTCTATTAATTTTAATCAATGAAGA	1499
Db	1437	GCATTACGATTAATATGATTTCAATCCAAATGATTTTCTATTAATTTTAATCAATGAAGA	1496
Qy	1500	GTTTTTCTATTAATGATTTAGTACACAGCATTAACAACAAGAACACCAACTCAGCAGCA	1559
Db	1497	GTTTTTCTATTAATGATTTAGTACACAGCATTAACAACAAGAACACCAACTCAGCAGCA	1556
Qy	1560	CACACACATTCAAATTCATTAATTAATGAGAGATGATGTTAAGTTAAGTTGAG	1619
Db	1557	CACACACATTCAAATTCATTAATTAATGAGAGATGATGTTAAGTTAAGTTGAG	1616
Qy	1620	AAAGTTCATTTCCAGGACATAATTTTGCCTCACACCACTCAGATAGAAAAACAACAAC	1679
Db	1617	AAAGTTCATTTCCAGGACATAATTTTGCCTCACACCACTCAGATAGAAAAACAACAAC	1676
Qy	1680	CTCATTTTGTATGTTTCCGAGTCCCTTGATTCGTTTTCCTTTTAAATCGAAATTAAG	1739
Db	1677	CTCATTTTGTATGTTTCCGAGTCCCTTGATTCGTTTTCCTTTTAAATCGAAATTAAG	1736
Qy	1740	CATATATGCATTAATTAATTAACAGATATACATTAACAATTAATTAATTAATTA	1799
Db	1737	CATATATGCATTAATTAATTAACAGATATACATTAACAATTAATTAATTAATTA	1796
Qy	1800	TGCTCTGATCTGTACATACATTTGTACAAAGTTTTTTTTAAAGCTAATTAATGCAAT	1859
Db	1797	TGCTCTGATCTGTACATACATTTGTACAAAGG --TTTTTAAAGCATTAATGCAAT	1854
Qy	1860	TTAATTAATGATCCGCAATCGAATTAACAATTTTGTTCATTTGATGAGTAAACAC	1919
Db	1855	TTAATTAATGATCCGCAATCGAATTAACAATTTTGTTCATTTGATGAGTAAACAC	1914
Qy	1920	TCTATCATGATGTAAAAAAGAAAAACAAGAAAAACATTTGATGATCTTTCGAAAAATA	1979
Db	1915	TCTATCATGATGTAAAAAAGAAAAACAAGAAAAACATTTGATGATCTTTCGAAAAATA	1974
Qy	1980	TATATCTAAACATAAAGAAAAACAACACTAGCTCTTAATAAATGCAAAAAATTT	2036
Db	1975	TATATCTAAACATAAAGAAAAACAACACTAGCTCTTAATAAATGCAAAAAATTT	2031

```

RESULT 2
US-10-108-605-116
; Sequence 116, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadue, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418

```

;; PRIOR FILING DATE: 2000-01-14
;; NUMBER OF SEQ ID NOS: 361
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 116
;; LENGTH: 2241
;; TYPE: DNA
;; ORGANISM: Drosophila melanogaster
US-10-108-605-116

Query Match 95.3%; Score 1940.4; DB 14; Length 2241;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2030; Conservative 0; Mismatches 6; Indels 56; Gaps 3;

QY 1 ATTCGTTGTCGACCGCTCCGAGCAAAAGATATGAAATCGAAGACGTAAACGAGAAAA 60
DB 89 ATTCGTTGTCGACCGCTCCGAGCAAAAGATATGAAATCGAAGACGTAAACGAGAAAA 148
QY 61 CACGCTGATTAAGTGTTCGTTTCGACATCGACAGAGTGAAGAAAGTGAAGTG 120
DB 149 CACGCTGATTAAGTGTTCGTTTCGACATCGACAGAGTGAAGAAAGTGAAGTG 208
QY 121 AGGCGCAAAAGCCGAAATCAATTAATGAAAAAGCCAGTCAAGAGCAATCTTTAACTGA 180
DB 209 AGGCGCAAAAGCCGAAATCAATTAATGAAAAAGCCAGTCAAGAGCAATCTTTAACTGA 268
QY 181 ACCTCGCAAGTAAATCCGGTACGACCCGTTAAACAACGCGCAAGATTCTACTCAAGAGCGCT 240
DB 269 ACCTCGCAAGTAAATCCGGTACGACCCGTTAAACAACGCGCAAGATTCTACTCAAGAGCGCT 328
QY 241 AAGCAATCAATCTCCAGAAATGAAGTCCCTGACGCGCTCTGCGACAGAGTGCTCCGGA 300
DB 329 AAGCAATCAATCTCCAGAAATGAAGTCCCTGACGCGCTCTGCGACAGAGTGCTCCGGA 388
QY 301 ATGCGCGCTTGAATGCGACGCGCGCGCATCCAGCGCATCCGACGACCGCGCGACGCGG 360
DB 389 ATGCGCGCTTGAATGCGACGCGCGCGCATCCAGCGCATCCGACGACCGCGCGACGCGG 448
QY 361 GAGACGCGGAGTGAAGATGTATCTGTCCAACTGAAGAACCTCGTTCCGTTCAATGCC 420
DB 449 GAGACGCGGAGTGAAGATGTATCTGTCCAACTGAAGAACCTCGTTCCGTTCAATGCC 508
QY 421 AAGACAGAAAGCTCAACAGCTGAGATCATCAGACGTCATCGATCATCTGCGAC 480
DB 509 AAGACAGAAAGCTCAACAGCTGAGATCATCAGACGTCATCGATCATCTGCGAC 568
QY 481 CTGCAAGCGAGCTGAGACGCAACCCGAGATGAGCACTTCGATGCGGACGCGCTCTG 540
DB 569 CTGCAAGCGAGCTGAGACGCAACCCGAGATGAGCACTTCGATGCGGACGCGCTCTG 628
QY 541 ACGGCGGTGAACGCACTCCACGAGACGAGACGAGCAAGCATGAGATGCGGATGCGGAG 600
DB 629 ACGGCGGTGAACGCACTCCACGAGACGAGACGAGCAAGCATGAGATGCGGATGCGGAG 668
QY 601 GCAGAAAGCGAAGTGCATCGATCATCTGCGCCAGCGCTGAAATGCGGACGCGGCG 660
DB 689 GCAGAAAGCGAAGTGCATCGATCATCTGCGCCAGCGCTGAAATGCGGACGCGGCG 748
QY 661 AAAAGTCTAAGTCCCGCGCGCTCTCCGCTTAACGATGCGCAACGCGCAACACTCTT 720
DB 749 AAAAGTCTAAGTCCCGCGCGCTCTCCGCTTAACGATGCGCAACGCGCAACACTCTT 808
QY 721 GTGCGCGCGCGCCATCCGAGACGATCAGACAGACAGACAGAACTGCAATGCGACGAA 780
DB 809 GTGCGCGCGCGCCATCCGAGACGATCAGACAGACAGACAGAACTGCAATGCGACGAA 868
QY 781 CAATGCAATCAGACGACAACTGTCCACAGATTAGCAAC-----821
DB 869 CAATGCAATCAGACGACAACTGTCCACAGATTAGCAACGCGTGAATTAAACAAT 928
QY 822 -----GCCACAGATGCGGAGAAAGACG 846
DB 929 TGCTAGTAATTTGAATCTTCGCTTCTCTTTTAGCACAAGATGCGGAGAAAGACG 988

QY 847 AGACAGTCTGAAGCGTGAAGAAAGCCGAGCGGAGAAACCCAGGAACTATAAGCTATT 906
DB 989 AGACAGTCTGAAGCGTGAAGAAAGCCGAGCGGAGAAACCCAGGAACTATAAGCTATT 1048
QY 907 AACTAATGATATGAAGAGTACAGTCAAACTCCACTGCGAACTTCTGTGAAGATCT 966
DB 1049 AACTAATGATATGAAGAGTACAGTCAAACTCCACTGCGAACTTCTGTGAAGATCT 1108
QY 967 CCGCATTAAGATTAACGCTGAACTCTCAAGAAACCCAGGCGGAGATCTCTTCAAC 1026
DB 1109 CCGCATTAAGATTAACGCTGAACTCTCAAGAAACCCAGGCGGAGATCTCTTCAAC 1168
QY 1027 GCTCCTTTTCCGCGGAGAAAGGAAACCCGCGGAGAAAGAGACCTGCTTTGCTCCACA 1086
DB 1169 GCTCCTTTTCCGCGGAGAAAGGAAACCCGCGGAGAAAGAGACCTGCTTTGCTCCACA 1228
QY 1087 GAGTTAGTTTAAACATTTAATACAGAAATACACACACAAACCTTAGCATTAAGCA 1146
DB 1229 GAGTTAGTTTAAACATTTAATACAGAAATACACACACAAACCTTAGCATTAAGCA 1288
QY 1147 GCGTAGCAAGTACATATATATTAATTAATTAACATTAACATTAAGATGAGCGCTT 1206
DB 1289 GCGTAGCAAGTACATATATATTAATTAATTAACATTAAGATGAGCGCTT 1348
QY 1207 AAGATGAACGAGAAAGAAAGAACTTGCAAGCTTTAAACAATTAATTTCTTTAAC 1266
DB 1349 AAGATGAACGAGAAAGAAAGAACTTGCAAGCTTTAAACAATTAATTTCTTTAAC 1408
QY 1267 GCTTGAATTAACATGAAGCATATATTTTGTTCCTAATTTAGTTTGTGTTT 1326
DB 1409 GCTTGAATTAACATGAAGCATATATTTTGTTCCTAATTTAGTTTGTGTTT 1468
QY 1327 AAAACAAGCAAGAAAGAAAGTATCTAATGCAAGCAAGCAAGCAAGCAAGCAAGCA 1386
DB 1469 AAAACAAGCAAGAAAGAAAGTATCTAATGCAAGCAAGCAAGCAAGCAAGCAAGCA 1528
QY 1387 ATATTAATTAATGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1446
DB 1529 ATATTAATTAATGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1588
QY 1447 GATTATATGATTTCAATCCAAATGATTTTCTAATTAATTTAATCAATGAAGATTTCC 1506
DB 1589 GATTATATGATTTCAATCCAAATGATTTTCTAATTAATTTAATCAATGAAGATTTCC 1648
QY 1507 TATTATGATTAAGCCACACCAATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1566
DB 1649 TATTATGATTAAGCCACACCAATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1708
QY 1567 ATTCAATTCATTAATTAATGAGAGATGAGTTAGATTTAGTTTGAAGAAAGTTC 1626
DB 1709 ATTCAATTCATTAATTAATGAGAGATGAGTTAGATTTAGTTTGAAGAAAGTTC 1768
QY 1627 ATTCCAGGACATTAATTTCACTCAACCACTCACTAATGAAGAAACCAACCACTTATT 1686
DB 1769 ATTCCAGGACATTAATTTCACTCAACCACTCACTAATGAAGAAACCAACCACTTATT 1828
QY 1687 TGTAGTTTCCAGATCCCTTGATGATGTTTCTCTTTTAAATCGAAATAGCATTAAG 1746
DB 1829 TGTAGTTTCCAGATCCCTTGATGATGTTTCTCTTTTAAATCGAAATAGCATTAAG 1888
QY 1747 CATTAATTAATTAACAGATATACCTAATACTAATTAATTAATTAATTAATTAATTA 1806
DB 1889 CATTAATTAATTAACAGATATACCTAATACTAATTAATTAATTAATTAATTAATTA 1948
QY 1807 TGAATGATGACATCTTTGTAACAAAGTGTTTTGTGATTAATTAATTAATTAATTA 1866
DB 1949 TGAATGATGACATCTTTGTAACAAAGTGTTTTGTGATTAATTAATTAATTAATTA 2008
QY 1867 TGAATGATGACATTAACATTTTGTGATTAATTAATTAATTAATTAATTAATTAATTA 1926
DB 2009 TGAATGATGACATTAATTAATTTTGTGATTAATTAATTAATTAATTAATTAATTA 2068
QY 1927 TGATGTAAAGAAAGAAAGCAAGTTC-ATGATGTTGCGAAAGAAATATTAAT 1985

Db 2069 TGAAGTAAAAAAAAAACAAGAAACATTGCATATGTGGAAAAAATATATTAT 2128
 Qy 1986 CTAAGACATAAGAAAAACAACACTAG-CTCTTATATTAATTCGAAAAATTT 2036
 Db 2129 CTAAGACATAAGAAAAACAACACTAGCTCTTATTAATTCGAAAAATTT 2180

RESULT 3

```

US-09-918-995-11565
: Sequence 11565, Application US/09918995
: Publication No. US20030073622A1
: GENERAL INFORMATION:
: APPLICANT: Hyeed, Inc.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: FROM INVENTION: FROM VARIOUS CDNA LIBRARIES
: FILE REFERENCE: 20411-756
: CURRENT APPLICATION NUMBER: US/09/918,995
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: US/09/235,076
: PRIOR FILING DATE: 1999-01-20
: NUMBER OF SEQ ID NOS: 38054
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 11565
: LENGTH: 481
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (1)..(481)
: OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11565

```

Query Match	4.1%;	Score 83.4;	DB 11;	Length 481;
Best Local Similarity	70.7%;	Pred. No. 5.7e-09;		
Matches 111; Conservative	0;	Mismatches 46;	Indels 0;	Gaps 0;

QY CGAGTGAAGAATATCTGTCCAACTGAAGACCTCGTTCGGTTCATGCCCAAGAAGC 428
Db CAACATGGAACGACTCTACTCCAACTCAAGAGGCTGAGGCCAGACATCCCCAAGAACA 308
QY GAAGTCAACCAAGCTGGAGATCATCAGCAGCTCATCGACTCATCTGGACCTGCAGAC 488
Db GAAGGTGAGCAAGATGGAATCTTGAGAGCATCTCATCACTACATCTTGGACCTGCAGAT 368
QY CGAGCTGAGACGCAACCCGAGATGGGCACTTCAT 525
Db CGCCCTGAGCTGCATCCCACTATTGTCAAGCCTGCAT 405

RESULT 4

```

US-09-919-580-165
; Sequence 165, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 165
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-165

```

Query Match	4.1%;	Score 83.4;	DB 10;	Length 530;
Best Local Similarity	70.7%;	Pred. No. 6.1e-09;		

	Matches	111; Conservative	0; Mismatches	46; Indels	0; Gaps
QY	369	CGAGATTAAGATGATCTGTCCAAACTGAAGAGACTGTCCGTTCAATGCCCAAGACAG	428		
Db	218	CAACATTAAGACTGTCTACTCAAGCTTCAAGAGACTGTCCAGACATCCCCAGAACAA	277		
QY	429	GAAGCTCACCAAGCTGAGATTCATCCAGACAGTATGATCTACATCTTCGACTTCAGAC	488		
Db	278	GAAGGTAGCAAGATGAGAAATCTCTGACAGACGTATGACTCTACATCTTGGACCTTCAGAT	337		
QY	489	CGAGCTTGAGACGACACCCCGAGATGGGCAACTTTCAT	525		
Db	338	CGCCCTTGACTCGCATCCCACTATTGTTCAGCTTCAT	374		

RESULT 5

```

US-09-919-580-228
Sequence 228, Application US/09919580
Patent No. US20020110832A1
GENERAL INFORMATION:
APPLICANT: Pyle, Ruth
APPLICANT: Xu, Jiangchun
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
OF COLON CANCER
FILE REFERENCE: 210121.552
CURRENT APPLICATION NUMBER: US/09/919,580
CURRENT FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 228
LENGTH: 532
TYPE: DNA
ORGANISM: Homo sapiens
US-09-919-580-228

```

Query Match	4.1%	Score 83.4	DB 10	Length 532
Best Local Similarity	70.7%	Pred. No. 6	1e-09	
Matches 111; Conservative	0	Mismatches 46	Indels 0	Gaps 0

Qy	Db	Qy	Db
365	CGAATATGAATGATATCTGTCCAACTGAAAGACCTGGTCCGTTCAAGCCCAAAACAG	428	
221	CAACATGACGACTCTACTCTCAAGGCTCAAGAGTGGTGGCCAGACATCCCAAGACAA	280	
429	GAAGCTCAACCAAGCGGAGATCATCCAGCAGTCAATGACTACATCTGCGAAGCTGCAGAC	488	
281	GAAAGTGGCAAGATGGAATCTCGCAGCAAGTCATCAATCAATCTTGGACCTGCAGAT	340	
489	CGAGCTGGAGACGACCCCGAGATGGGCAATTCGAT	525	
341	CGCCCTGGACTCGATCCCACTATTGTGAGCTCGAT	377	

RESULT 6

```

US-09-919-580-538
? Sequence 538: Application US/09919580
? Patent No. US20020110832A1
? GENERAL INFORMATION:
? APPLICANT: Pyle, Ruth
? APPLICANT: Xu, Jiangchun
? APPLICANT: Secrist, Heather
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
? TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
? FILE REFERENCE: 210121.552
? CURRENT APPLICATION NUMBER: US/09/919,580
? CURRENT FILING DATE: 2001-07-30
? NUMBER OF SEQ ID NOS: 934
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 538
? LENGTH: 533
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
/

```


NAME/KEY: misc.feature
LOCATION: 7, 8, 9, 524
OTHER INFORMATION: n = A,T,C or G
US-09-919-580-538

Query Match 4.1%; Score 83.4; DB 10; Length 533;
Best Local Similarity 70.7%; Pred. No. 6.1e-09;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATGTATCTGTCCAACTGAAGAGACCTGTCCTGATGCTCCCAAGAACAG 428
DB 221 CAACATGAAGAGCTGTACTTCAAGCTCAAGAGAGCTGTGCTCCAGCATCCCAAGAACAA 280
QY 429 GAAGCTCACCAAGCTGAGATCATCCAGCATGTCATGATCATCTGCAAGCTGCAGAC 488
DB 281 GAAGGTGAGCAAGATGGAATCTCTGACGACGTCATGATCATCTTGGAACTGCAGAT 340
QY 489 CGAGCTGAGAGACGACCCCGAGATGGGCAACTTCGAT 525
DB 341 CGCCCTGAGACTCGCATCCCACTATTGTGACCTGCAT 377

RESULT 7
US-10-096-534-30

Sequence 30, Application US/10096534
Publication No. US2003016887A1

GENERAL INFORMATION:

APPLICANT: The Brigham and Women's Hospital, Inc.

APPLICANT: Yates, Karen

APPLICANT: Mizuno, Shuichi

APPLICANT: Glowacki, Julie

TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF SKELETAL DEGENERATION CONDITIONS

FILE REFERENCE: B0801/7244/KA/RP

CURRENT APPLICATION NUMBER: US/10/096,534

CURRENT FILING DATE: 2002-03-12

PRIOR APPLICATION NUMBER: US 60/274,980

PRIOR FILING DATE: 2001-03-12

NUMBER OF SEQ ID NOS: 79

SOFTWARE: PatentIn version 3.0

SEQ ID NO 30

LENGTH: 1049

TYPE: DNA

ORGANISM: Homo sapiens

US-10-096-534-30

Query Match 4.1%; Score 83.4; DB 13; Length 1049;
Best Local Similarity 70.7%; Pred. No. 9.8e-09;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATGTATCTGTCCAACTGAAGAGACCTGTCCTGATGCTCCCAAGAACAG 428
DB 207 CAACATGAAGAGCTGTACTTCAAGCTCAAGAGAGCTGTGCTCCAGCATCCCAAGAACAA 266
QY 429 GAAGCTCACCAAGCTGAGATCATCCAGCATGTCATGATCATCTGCAAGCTGCAGAC 488
DB 267 GAAGGTGAGCAAGATGGAATCTCTGACGACGTCATGATCATCTTGGAACTGCAGAT 326
QY 489 CGAGCTGAGAGACGACCCCGAGATGGGCAACTTCGAT 525
DB 327 CGCCCTGAGACTCGCATCCCACTATTGTGACCTGCAT 363

RESULT 8

US-09-925-300-607

Sequence 607, Application US/09925300

Patent No. US20020151681A1

GENERAL INFORMATION:

APPLICANT: Craig Rosen,

APPLICANT: Steve Ruden,

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA101

CURRENT APPLICATION NUMBER: US/09/925,300

CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05988

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 1890

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 607

LENGTH: 1348

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1328)

OTHER INFORMATION: n equals a,t,g, or c

US-09-925-300-607

Query Match 4.1%; Score 83.4; DB 10; Length 1348;
Best Local Similarity 70.7%; Pred. No. 1.2e-08;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATGTATCTGTCCAACTGAAGAGACCTGTCCTGATGCTCCCAAGAACAG 428
DB 237 CAACATGAAGAGCTGTACTTCAAGCTCAAGAGAGCTGTGCTCCAGCATCCCAAGAACAA 296
QY 429 GAAGCTCACCAAGCTGAGATCATCCAGCATGTCATGATCATCTGCAAGCTGCAGAC 488
DB 297 GAAGGTGAGCAAGATGGAATCTCTGACGACGTCATGATCATCTTGGAACTGCAGAT 356
QY 489 CGAGCTGAGAGACGACCCCGAGATGGGCAACTTCGAT 525
DB 357 CGCCCTGAGACTCGCATCCCACTATTGTGACCTGCAT 393

RESULT 9

US-10-115-831-69

Sequence 69, Application US/10115831

Publication No. US20030219743A1

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Ren, Feiyun

APPLICANT: Dmanac, Radoje T.

TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and

FILE REFERENCE: Polypeptides

CURRENT APPLICATION NUMBER: US/10/115,831

CURRENT FILING DATE: 2002-04-02

PRIOR APPLICATION NUMBER: 09/667,298

PRIOR FILING DATE: 2000-09-22

PRIOR APPLICATION NUMBER: 09/577,408

PRIOR FILING DATE: 2000-05-18

NUMBER OF SEQ ID NOS: 178

SOFTWARE: PC FL_genes Version 2.0

SEQ ID NO 69

LENGTH: 1397

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (52)..(612)

US-10-115-831-69

Query Match 4.1%; Score 83.4; DB 13; Length 1397;
Best Local Similarity 70.7%; Pred. No. 1.2e-08;
Matches 111; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAAGATGTATCTGTCCAACTGAAGAGACCTGTCCTGATGCTCCCAAGAACAG 428
DB 318 CAACATGAAGAGCTGTACTTCAAGCTCAAGAGAGCTGTGCTCCAGCATCCCAAGAACAA 377
QY 429 GAAGCTCACCAAGCTGAGATCATCCAGCATGTCATGATCATCTTGGAACTGCAGAT 488

Db 378 GAAGTGAAGAGATGAAATCTGACGACGTATCGATCTTGAGCCTGCAGAT 437
QY 489 CGAGCTGAGACGACCCCGAGATGGCACTTCGAT 525
Db 438 CGCCTGGACTCGCATCCACTATTGTGACCTGCAT 474

RESULT 10

US-09-764-872-494/c
; Sequence 494, Application US/09764872
; Publication No. US2003050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 494
; LENGTH: 2511
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-872-494

Query Match 4.1%; Score 83.4; DB 11; Length 2511;
Best Local Similarity 70.7%; Pred. No. 1.8e-08;
Matches 11; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 369 CGAGATGAGATGATCTGTCCAAATGAGACCTGTTCCGTTCAATGCCAAGAC 428
Db 2179 CACATGAGACGACTGCTACTCCAGCTCAGAGACTGTGCCAGCATCCCCAGAACAA 2120
QY 429 GAAGCTGACCAAGCTGAGATCATCCAGCATGCTGACTCATCTGCGAAGCTGCAGAC 488
Db 2119 GAAAGTGAGCAAGATGAAATCTCGACGACGTCATCGATCACTTGGACCTGCAGAT 2060
QY 489 CGAGCTGAGACGACCCCGAGATGGCACTTCGAT 525
Db 2059 CGCCTGGACTCGCATCCACTATTGTGACCTGCAT 2023

RESULT 11

US-10-025-170-1
; Sequence 1, Application US/10025170
; Publication No. US20030003467A1
; GENERAL INFORMATION:
; APPLICANT: Albert Einstein College of Medicine of Yeshiva University
; APPLICANT: Iavarone, Antonio
; APPLICANT: Laboralle, Anna
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND TREATING PEDIATRIC NEOPLASMS
; FILE REFERENCE: 96700/709
; CURRENT APPLICATION NUMBER: US/10/025,170
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/257,847
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-025-170-1

Query Match 4.0%; Score 81.8; DB 15; Length 405;
Best Local Similarity 70.1%; Pred. No. 1.2e-08;
Matches 110; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 369 CGAGATGAGATGATCTGTCCAAATGAGACCTGTTCCGTTCAATGCCAAGAC 428
Db 111 CACATGAGACGACTGCTACTCCAGCTCAGAGACTGTGCCAGCATCCCCAGAACAA 170

QY 429 GAAGCTACCAAGCTGAGATCATCCAGACGTCATCGATCATCTGTGAGCCTGCAGAC 488
Db 171 GAAGTGAGAGAGATGAAATCTGACGACCTCATCGATCATCTTGAGCCTGCAGAT 230
QY 489 CGAGCTGAGACGACCCCGAGATGGCACTTCGAT 525
Db 231 CGCCTGGACTCGCATCCACTATTGTGACCTGCAT 267

RESULT 12

US-10-291-172-44
; Sequence 44, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 44
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)..(548)
US-10-291-172-44

Query Match 3.9%; Score 80; DB 12; Length 979;
Best Local Similarity 58.5%; Pred. No. 5.7e-08;
Matches 159; Conservative 0; Mismatches 110; Indels 3; Gaps 1;

QY 345 GCACCGCGGGAAGGGGAGAACCGCGAGATGATGTCTCCAACTGAGGACCT 404
Db 260 GCAGCAGGTAAAGTGTCTCTTACGACATGAACGGCTGTACTCAGCCTCAAGAGCT 319
QY 405 CGTTCGTTATGCTCCCAAGACAGAAAGCTCACCAAGCTGAGATCATCCAGCAGTCAT 464
Db 320 GTGTCCACCTCTGCTCCCAAGACCGAAAGTGAAGATGATCTCCAGCAGTCAT 379
QY 465 CCACTACATGTGAGACTGAGACCGAAGCTGAGACGACCCCGAGATGGCACTTCGA 524
Db 380 CGACTACATCAGGAGACTTCAAGTGAAGCTGAGATCCGAAGTTGAAACCCCGG 439
QY 525 TGGGGCAG--CCGCTTGAAGCGGCTGAACCGAATCTCCAGAGAGAGAGAGCAT 581
Db 440 GGGCCGAGGGGCTCGGCTCGGCTCGCTCAACGACCTCAACGGAGATCAGCGCCT 499
QY 582 GGAGATGCGGATGCCGAGGAGAGAGGAG 613
Db 500 GACGCGCAGGCGGAGATGCTTCTTCCGAGCG 531

RESULT 13

US-10-119-428-47
; Sequence 47, Application US/10119428
; Publication No. US20030165881A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua

```
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Wehrman, Tom
APPLICANT: Ren, Feiyan
APPLICANT: Ma, Yundqing
APPLICANT: Zhou, Ping
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonghong
APPLICANT: Dymnac, Radoje T.
TITLE OF INVENTION: No. US20030165881A1el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/10/119,428
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/596,193
PRIOR FILING DATE: 2000-06-17
PRIOR APPLICATION NUMBER: 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519,705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 55
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 47
LENGTH: 1162
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (267)..(734)
NAME/KEY: misc feature
LOCATION: (1)..(1162)
OTHER INFORMATION: n = a,t,c or g
US-10-119-428-47
```

Query Match 3.9%; Score 80; DB 13; Length 1162;
Best Local Similarity 58.5%; Pred. No. 6.5e-08;
Matches 159; Conservative 0; Mismatches 110; Indels 3; Gaps 1;

```
QY 345 GCACCGCGGCGGAGAGACCGCGAGATGATGATCTCTCCAACTGAAGACCT 404
DB 443 GCAGCAGGTAAAGTGTCTCTACGACATGAACGGCTGTACTCAGCTCAAGGAGCT 502
QY 405 CGTTCGTTCAATCCCAAGAACGAGATCACCAGCTGAGATCATCTCCAGACGCTAT 464
DB 503 GGTGCCACCTCCCGACGACGCAAGGTGAGCAAGTGAATCTCCAGACGCTCAT 562
QY 465 CGACTACATCTGCGACCTGCGACACCGAGCTGAGAGCACCAGAGATGGCACTTGA 524
DB 563 CGACTACATCAGGAGACCTTCACTGAGCTGAACTCGGAATCCGAAGTTGAAACCCCGG 622
QY 525 TGGCGGAG---CCGCTCTGACGGCGGTGAACGGAATCCACAGAGACGAGACGACAT 581
DB 623 GGGCCGAGGGCTGCTCCGGCTCGGGCTCGGCTCAGACACCTCAACGGCGAGTCAAGCGCCT 682
QY 582 GGAGATGCGGATCCGAGCGAGCAAGAGCGGAG 613
DB 683 GACGCGCGAGGCGGACATGCTTCCTCGGAGCG 714
```

RESULT 14
US-10-027-632-141641

```
Sequence 141641, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
```

```
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 141641
LENGTH: 663
TYPE: DNA
ORGANISM: Human
US-10-027-632-141641
```

Query Match 3.9%; Score 79.8; DB 13; Length 663;
Best Local Similarity 65.4%; Pred. No. 4.9e-08;
Matches 117; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
QY 345 GCACCGCGGCGGAGAGACCGCGAGATGATGATCTCTCCAACTGAAGACCT 404
DB 129 GCAGCAGGTAAAGTGTCTCTACGACATGAACGGCTGTACTCAGCTCAAGAGCT 188
QY 405 CGTTCGTTCAATCCCAAGAACGAGATCACCAGCTGAGATCATCTCCAGACGCTAT 464
DB 189 GGTGCCACCTCCCGACGACGCAAGGTGAGCAAGTGAATCTCCAGACGCTCAT 248
QY 465 CGACTACATCTGCGACCTGCGACACCGAGCTGAGAGCACCAGATGGCACTTGG 523
DB 249 CGACTACATCAGGAGACCTTCAGTTGAGCTGAATCTGGAATCCGAATGGACCCCG 307
```

RESULT 15
US-10-027-632-141641

```
Sequence 141641, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 141641
LENGTH: 663
TYPE: DNA
ORGANISM: Human
US-10-027-632-141641
```

Query Match 3.9%; Score 79.8; DB 14; Length 663;
Best Local Similarity 65.4%; Pred. No. 4.9e-08;
Matches 117; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Mon Feb 9 08:28:26 2004

us-09-614-150a-11.rnpb

Page 8

QY 345 GCACCGCGCGACGGGGAGAAAGCCGAGATGAAGATGTATCTGTCCAAATGAAGGACT 404

Db 129 GGAGCAGGTAAACGTGTCTCTTACACATGAAGCGCTGTACTACGCTCAAGAGCT 188

QY 405 CGTTCCGTTTCATGCCAAGAACGAAAGCTCACCAAGCTTGAGATATCCAGACGTCAT 464

Db 189 GGTGCCACCCCTGTCCCAAGAACGCAAGGTGAGCAAGTGGAGATTCTCCAGACGTTCAT 248

QY 465 CGACTTACATCTGCGACCTTGACAGCCGAGCTTGAGAGCGCACCCCGAGATGGGCAACTTGG 523

Db 249 CGACTTACATCAGGAGACTTTCAGTTGGAGGTGAATCTGGAATCCGAGTTGGATCCCGG 307

Search completed: February 8, 2004, 20:44:48
Job time : 774.996 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2004, 18:36:25 ; Search time 4775.14 Seconds
(without alignments)
10367.824 Million cell updates/sec

Title: US-09-614-150A-11

Perfect score: 2036
Sequence: 1 attcgtctgcctgcgcctcg.....ttaataatgcacaaattc 2036

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estda:*
2: em_esthm:*
3: em_estin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phy:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	684	33.6	684	12	BI238272 RE34679.5
2	673	33.1	675	12	BI367923 RE53804.5
3	673	33.1	677	12	BI367945 RE53827.5
4	670.4	32.9	672	12	BI234348 RE30030.5

5	665	32.7	669	12	BI234976	BI234976 RE30759.5
6	664.4	32.6	667	12	BI613765	BI613765 RH42996.5
7	661.6	32.5	668	9	A1134649	A1134649 GH12170.5
8	644.8	31.7	651	10	BG636933	BG636933 SD14761.5
9	626	30.7	631	12	BI655419	BI655419 RH63190.5
10	625.4	30.7	631	12	BI227773	BI227773 RE24996.5
11	624	30.6	628	12	BI232704	BI232704 RE28821.5
12	622.4	30.6	635	12	BI637096	BI637096 SD19014.5
13	620.4	30.5	622	12	BI485228	BI485228 RE68523.5
14	619	30.4	623	12	BI363964	BI363964 RA48965.5
15	618.4	30.4	620	12	BI173488	BI173488 RH16419.5
16	614.6	30.2	621	12	BI374484	BI374484 RE62085.5
17	614	30.2	626	12	BI216272	BI216272 RE23251.5
18	613.8	30.1	622	9	AA567404	AA567404 HL01136.5
19	602	29.6	606	12	BI370607	BI370607 RE57307.5
20	598.8	29.4	603	12	BI575395	BI575395 RH31018.5
21	598	29.4	598	12	BI61601	BI61601 RE45887.5
22	596.4	29.3	598	12	BI232788	BI232788 RE28918.5
23	592.2	29.1	630	9	AA802644	AA802644 GM06052.5
24	574.4	28.2	598	9	AA391523	AA391523 LD10532.5
25	566	27.8	566	12	BI364246	BI364246 RE49317.5
26	562	27.6	562	12	BI365829	BI365829 RE52117.5
27	560.4	27.5	562	12	BI36776	BI36776 RE32890.5
28	551.8	27.1	562	10	AW942140	AW942140 LD10532.3
29	546	26.8	547	12	BI566493	BI566493 RH36054.5
30	543.2	26.7	561	12	BI584422	BI584422 RH23621.5
31	543	26.7	548	12	BI582179	BI582179 RH20131.5
32	542	26.6	546	12	BI173908	BI173908 RE17045.5
33	541	26.6	543	12	BI593547	BI593547 RH11647.5
34	540.4	26.5	543	12	BI624214	BI624214 RH63736.5
35	540	26.5	542	12	BI608402	BI608402 RH12893.5
36	532	26.1	536	12	BI234067	BI234067 RE29703.5
37	525.2	25.8	545	12	BI481547	BI481547 RE64235.5
38	522	25.6	538	12	BI175978	BI175978 RE63893.5
39	483.6	23.8	545	12	BI169977	BI169977 RE11306.5
40	483	23.7	485	12	BI588404	BI588404 RH29609.5
41	459	22.5	460	12	BI615553	BI615553 RH45229.5
42	454.4	22.3	489	9	AA439247	AA439247 LD13757.5
43	439.6	21.6	517	29	AG215595	AG215595 Drosophila
44	403	19.8	409	12	BI65521	BI65521 RE50865.5
45	346	17.0	529	29	AG215543	AG215543 Drosophila

ALIGNMENTS

RESULT 1
BI238272
LOCUS
DEFINITION RE34679.5prtime RE Drosophila melanogaster normalizd Embryo pRLc-1
Drosophila melanogaster cDNA clone RE34679 5 similar to emc:
Fban001007 'transcription factor' located on: 3L 61D1-61D2;
05/12/2001, mRNA sequence.

ACCESSION BI238272.1 GI:14706766
VERSION BI238272
KEYWORDS

SOURCE EST.
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Bphnoidae; Drosophilidae; Drosophila.
1 (bases 1 to 684)

AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Friese,E., George
R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S.,
Mungall,C.J., Nuno,J., Pacle,J., Paragas,V., Park,S.,
Phonnanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
G.M.

TITLE BDGP/HMI RE Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab

QY 183 CTCGACGATATCCGCTGACGACCGGTGTAACACCGCAAGATTCTACTCAAGAGCGTAA 242
DB 183 CTCGACGATATCCGCTGACGACCGGTGTAACACCGCAAGATTCTACTCAAGAGCGTAA 242
QY 243 AGCAATATCTCCGAATGAAGTCCCTGACGCGCTCTGCGACAGAGTGTCTCCGAAT 302
DB 243 AGCAATATCTCCGAATGAAGTCCCTGACGCGCTCTGCGACAGAGTGTCTCCGAAT 302
QY 303 GCCGCGCTTGAATCCGACGCGCGCATCCAGCGCATCCGACGCGCGCGCGCGGGA 362
DB 303 GCCGCGCTTGAATCCGACGCGCGCATCCAGCGCATCCGACGCGCGCGCGCGGGA 362
QY 363 GAAAGCGGAGATGAATGTAATCTGTCCAACTGAGAGAGCTCTGCTGATGCCAA 422
DB 363 GAAAGCGGAGATGAATGTAATCTGTCCAACTGAGAGAGCTCTGCTGATGCCAA 422
QY 423 GAACAGAGATCTCCAGCTGAGATCTCCAGAGATCTCCAGAGATCTCCAGAGATCT 482
DB 423 GAACAGAGATCTCCAGCTGAGATCTCCAGAGATCTCCAGAGATCTCCAGAGATCT 482
QY 483 GCAGACCGAGCTGAGAGCGACCGCGGATGCGCACTTGCATGCGGCGCGCTGAC 542
DB 483 GCAGACCGAGCTGAGAGCGACCGCGGATGCGCACTTGCATGCGGCGCGCTGAC 542
QY 543 GCGGCTGAACGAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 602
DB 543 GCGGCTGAACGAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 602
QY 603 AGAAGCGGAGATGATCTCAATATCTCTGCGCGCGCTGAGATCCGAGAGAGAGAG 662
DB 603 AGAAGCGGAGATGATCTCAATATCTCTGCGCGCGCTGAGATCCGAGAGAGAGAG 662
QY 663 AGTCTCTAGTCCC 675
DB 663 AGTCTCTAGTCCC 675

RESULT 3
BI367945 677 bp mRNA linear EST 01-AUG-2001
LOCUS RS53827.5prime RE Drosophila melanogaster normalized Embryo pFLC-1
DEFINITION Drosophila melanogaster cDNA clone RS53827 5 similar to emc:
Pfam001007 'transcription factor' located on: 3L 61D1-61D2;:
05/14/2001, mRNA sequence.

ACCESSION BI367945
VERSION BI367945.1 GI:15063973
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Phyloglia; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 677)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
R., Gonzalez, M., Guarini, H., Harris, N., Li, P., Liao, G., Miera, S.,
Mungall, C.J., Nuno, J., Pacle, J., Pargade, V., Park, S.,
Phouenavong, S., Wan, K., Yu, C., Lewis, S.B., Celisner, S. and Rubin
G.M.
TITLE BDGP/HMM RE Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AB003469: arm:3L 608641, 911167
estimated-cyto:61C8-61B1: 05/14/2001
Plate: RE:538 row: C column: 3
High quality sequence stop: 623.
Location/Qualifiers
1. 677

/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE53827"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFLC-1"
note="Organ: embryo; Vector: pFLC1; Site: 1; XhoI; Site: 2;
BamHI. Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
BASE COUNT 199 a 188 c 181 g 109 t
ORIGIN
Query Match 33.1%; Score 673; DB 12; Length 677;
Best Local Similarity 100.0%; Pred. No. 2.6e-77;
Matches 673; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATTGTTGCTGACCGCTGACGACAAAGATATGAAGATGAAGATGAAGATGAAGATGA 60
DB 5 ATTGTTGCTGACCGCTGACGACAAAGATATGAAGATGAAGATGAAGATGAAGATGA 64
QY 61 CAGCTGATTAAGTGTTCGTTTCGAAATGATGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 65 CAGCTGATTAAGTGTTCGTTTCGAAATGATGAGAGAGAGAGAGAGAGAGAGAGAG 124
QY 121 AGGCGACAAAGCCGAATCAATCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 125 AGGCGACAAAGCCGAATCAATCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 184
QY 181 ACTGCGAAGTATCCGCTGACGACCGGTGTAACACCGCAAGATTCTACTCAAGAGAG 240
DB 185 ACTGCGAAGTATCCGCTGACGACCGGTGTAACACCGCAAGATTCTACTCAAGAGAG 244
QY 241 AAAGCAATATCTCCGAATGAAGTCCCTGACGCGCGCTGACAGAGAGAGAGAGAG 300
DB 245 AAAGCAATATCTCCGAATGAAGTCCCTGACGCGCGCTGACAGAGAGAGAGAGAG 304
QY 301 ATGCGCGCTTGAATCCAGCGCGCGCATCCAGCGCGCATCCAGCGCGCGCGCGAG 360
DB 305 ATGCGCGCTTGAATCCAGCGCGCGCATCCAGCGCGCATCCAGCGCGCGCGCGAG 364
QY 361 GAGAGCGCGAGATGAAGTGTATCTGTCCAACTGAAGAGAGAGAGAGAGAGAGAGAG 420
DB 365 GAGAGCGCGAGATGAAGTGTATCTGTCCAACTGAAGAGAGAGAGAGAGAGAGAGAG 424
QY 421 AAGACAGAGAGATCCAGAGCTGAGATCATCCAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 425 AAGACAGAGAGATCCAGAGCTGAGATCATCCAGAGAGAGAGAGAGAGAGAGAGAG 484
QY 481 CTCGACGAGAGCTGAG 540
DB 485 CTCGACGAGAGCTGAG 544
QY 541 ACGCGCGTGAACGAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 545 ACGCGCGTGAACGAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
QY 601 GCAGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 605 GCAGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664
QY 661 AAAGTCTTAGTC 673
DB 665 AAAGTCTTAGTC 677

RESULT 4
BI234348 672 bp mRNA linear EST 12-JUL-2001
LOCUS RS30030.5prime RE Drosophila melanogaster normalized Embryo pFLC-1
DEFINITION

Drosophila melanogaster cDNA clone RE30030 5 similar to emc:
Phan001007, 'transcription factor' located on: 3L 61D1-61D2;:
04/12/2001, mRNA sequence.

ACCESSION BI234348.1 GI:14702783

VERSION BI234348.1

KEYWORDS EST.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

REFERENCE Eukaryota: Metazoa: Arthropoda: Hexapoda: Insecta: Pterygota;
Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha;
Ephydroidea: Drosophilidae: Drosophila.

AUTHORS 1 (bases 1 to 672)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
, J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
, R., Gonzalez, M., Guartin, H., Harris, N., Li, P., Liao, G., Misra, S.,
Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin
, G.M.

TITLE BDGP/HHMI RE Drosophila EST Project

JOURNAL Unpublished

COMMENT Contact: Stapleton, M.

BDGP Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic: AB003469: arm:3L [608641,911167]
estimated-cyto:61C8-61E1: 04/12/2001
Plate: RE.300 row: C column: 6
High quality sequence stop: 615.
Location/Qualifiers

FEATURES

Source 1..672
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE30030"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pfic-1"
/note="Organ: embryo; Vector: pFic1; Site 1: XhoI; Site 2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."

BASE COUNT 199 a 184 c 181 g 108 t

ORIGIN

Query Match 32.9%; Score 670.4; DB 12; Length 672;
Best Local Similarity 99.9%; Pred. No. 5.5e-77;
Matches 671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTCGTTGTGACCGCTCGCAGCAAAAGATATGAAATCGAAGACGTAACGAGAAA 60
DB 1 ATTCGTTGTGACCGCTCGCAGCAAAAGATATGAAATCGAAGACGTAACGAGAAA 60

QY 61 CACCTGATTAAGTGTGTTGTTTCCAACTCGACAGAGAGAGAGAGAGAGAGAGAG 120
DB 61 CACCTGATTAAGTGTGTTGTTTCCAACTCGACAGAGAGAGAGAGAGAGAGAGAG 120

QY 121 AGCCGACAAAGCCGAAATCAATCAAAATTTGAAAAAGCCAGTCAAGAGCATCTTAACTGA 180
DB 121 AGCCGACAAAGCCGAAATCAATCAAAATTTGAAAAAGCCAGTCAAGAGCATCTTAACTGA 180

QY 121 AGCCGACAAAGCCGAAATCAATCAAAATTTGAAAAAGCCAGTCAAGAGCATCTTAACTGA 180
DB 121 AGCCGACAAAGCCGAAATCAATCAAAATTTGAAAAAGCCAGTCAAGAGCATCTTAACTGA 180

QY 181 ACCTCGCAAGTAATCCGCTAGACCGGTGTAACAACCGCAAGATTCTTCAAGAGAGGT 240
DB 181 ACCTCGCAAGTAATCCGCTAGACCGGTGTAACAACCGCAAGATTCTTCAAGAGAGGT 240

QY 241 AAAACATCATCTCCAGAAATGAATGCTCCGAGCCGCTGTCGAGACAGAGTCTCCGGA 300
DB 241 AAAACATCATCTCCAGAAATGAATGCTCCGAGCCGCTGTCGAGACAGAGTCTCCGGA 300

QY 301 ATGCCGCGCTGATGCGACGGGCGCATTCAGCGCATCCAGCAGCAACCGGCGAGCGGG 360

DB 301 ATGCCGCGCTGATGCGACGGGCGCATTCAGCGCATCCAGCAGCAACCGGCGAGCGGG 360

QY 361 GAGACGCGGAGATGAAATGATCTGTCCAAATCTGAAGAGACTCGTTCGTCATGCCC 420
DB 361 GAGACGCGGAGATGAAATGATCTGTCCAAATCTGAAGAGACTCGTTCGTCATGCCC 420

QY 421 AAGACAGAAAGCTCAACCAAGTGAATCATCAGACGTCATGCAATCATCTGCGAC 480
DB 421 AAGACAGAAAGCTCAACCAAGTGAATCATCAGACGTCATGCAATCATCTGCGAC 480

QY 481 CTGCAAGCAGAGCTCGAGACGACCCCGAGATGGGCACTTGAATGCGAGCCGCTCTG 540
DB 481 CTGCAAGCAGAGCTCGAGACGACCCCGAGATGGGCACTTGAATGCGAGCCGCTCTG 540

QY 541 ACAGCGGTGAAGCAGCTCCACGAGACGAGACAGAGCACTGAGAGATGCGAGTCCGAG 600
DB 541 ACAGCGGTGAAGCAGCTCCACGAGACGAGACAGAGCACTGAGAGATGCGAGTCCGAG 600

QY 601 GCAGAAAGCGAAGTCAATCCAGATCTCTGCGCCAGCGCTGAATGCCAGACCGGCG 660
DB 601 GCAGAAAGCGAAGTCAATCCAGATCTCTGCGCCAGCGCTGAATGCCAGACCGGCG 660

QY 661 AAAGTCTTAGT 672
DB 661 AAAGTCTTAGT 672

RESULT 5 669 bp mRNA linear EST 12-JUL-2001
BI234976 RE30759.5prime RE Drosophila melanogaster normalized Embryo pFic-1
LOCUS Drosophila melanogaster cDNA clone RE30759 5 similar to emc:
Phan001007, 'transcription factor' located on: 3L 61D1-61D2;:
05/12/2001, mRNA sequence.

ACCESSION BI234976 GI:14703419

VERSION BI234976

KEYWORDS EST.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

REFERENCE Eukaryota: Metazoa: Arthropoda: Hexapoda: Insecta: Pterygota;
Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha;
Ephydroidea: Drosophilidae: Drosophila.

AUTHORS 1 (bases 1 to 669)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
, J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
, R., Gonzalez, M., Guartin, H., Harris, N., Li, P., Liao, G., Misra, S.,
Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin
, G.M.

TITLE BDGP/HHMI RE Drosophila EST Project

JOURNAL Unpublished

COMMENT Contact: Stapleton, M.

BDGP Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic: AB003469: arm:3L [608641,911167]
estimated-cyto:61C8-61E1: 05/12/2001
Plate: RE.307 row: E column: 11
High quality sequence stop: 583.
Location/Qualifiers

FEATURES

Source 1..669
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE30759"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pfic-1"

/note="Organ: embryo; Vector: pFic1; Site 1: XhoI; Site 2: BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

BASE COUNT 198 a 185 c 180 g 106 t

Query Match 32.7%; Score 665; DB 12; Length 669;
Best Local Similarity 100.0%; Pred. No. 2.7e-76;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATTGTTGTCGACCGCTCGACGAAAGATATGAAATCGAGACGTAACGCAAAA 60
DB 5 ATTGTTGTCGACCGCTCGACGAAAGATATGAAATCGAGACGTAACGCAAAA 64
QY 61 CAGCTGGATTAAAGTTCCTTCGTAACGACGAGTGAAGAAAACAAAGTGAAGT 120
DB 65 CAGCTGGATTAAAGTTCCTTCGTAACGACGAGTGAAGAAAACAAAGTGAAGT 124
QY 121 AGGCAACAAAGCCGAATCAATCAATTTGAAAGCCAGTCAAGAGCAATCTTTA 180
DB 125 AGGCAACAAAGCCGAATCAATCAATTTGAAAGCCAGTCAAGAGCAATCTTTA 184
QY 181 ACCTCGCAAGTAATCCGTAACGACCGTGTAAACAAACGCAAGATTCTTCAAG 240
DB 185 ACCTCGCAAGTAATCCGTAACGACCGTGTAAACAAACGCAAGATTCTTCAAG 244
QY 241 AAAGCAATCATCTTCGAAATGAAGTCCCTGACGCGCTTCGCAAGAGTCCCT 300
DB 245 AAAGCAATCATCTTCGAAATGAAGTCCCTGACGCGCTTCGCAAGAGTCCCT 304
QY 301 ATGCGCGCTTGAATGCGAGCGGCGCATCGAGCGCCATCCAGCAACGCGGGA 360
DB 305 ATGCGCGCTTGAATGCGAGCGGCGCATCGAGCGCCATCCAGCAACGCGGGA 364
QY 361 GAGAACGCGCGAGATGAAGATGTATCTGTCCAACTGAAGAACCTGTTCCGTT 420
DB 365 GAGAACGCGCGAGATGAAGATGTATCTGTCCAACTGAAGAACCTGTTCCGTT 424
QY 421 AAAGACGAAAGCTCACCAAGCTGAGATCATTCAGACGTCATCATCTTGAC 480
DB 425 AAAGACGAAAGCTCACCAAGCTGAGATCATTCAGACGTCATCATCTTGAC 484
QY 481 CTGCAAGCGAGCTGAGAGCGACCCGAGATGCGGCACTTCATGCGGAGCGCT 540
DB 485 CTGCAAGCGAGCTGAGAGCGACCCGAGATGCGGCACTTCATGCGGAGCGCT 544
QY 541 ACCGCGGTGAACGACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 545 ACCGCGGTGAACGACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
QY 601 GCGAAGCGGAAGTGCATCCAGATATCTTCGCGCCAGCGCTGAATGCGAG 660
DB 605 GCGAAGCGGAAGTGCATCCAGATATCTTCGCGCCAGCGCTGAATGCGAG 664
QY 661 AAAGT 665
DB 665 AAAGT 669

```

RESULT 6
LOCUS B1613765 667 bp mRNA linear EST 07-SEP-2001
DEFINITION RH42996.5prime RH Drosophila melanogaster normalized Head pFic-1
Drosophila melanogaster cDNA clone RH42996 5 similar to emc:
Rhan0001007 GO: [nucleus (GO:0005634); transcription co-repressor
(GO:0003714)] located on: 3L 61D1-61D2;: 08/18/2001, mRNA sequence.

ACCESSION B1613765
VERSION B1613765.1 GI:15509290
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 667)

REFERENCE
AUTHORS Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mira, S.,
Mungall, C. J., Nuno, J., Pacle, J., Parag, V., Park, S.,
Phouenavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and Rubin
G. M.
BDGP/HMI RH Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP

TITLE
JOURNAL
COMMENT
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
hit genomic AB003469: arm:3L [608641,91167]
estimated-cyto:61C8-61E1: 08/18/2001
plate: RH 429 row: H column: 12
High quality sequence stop: 619.
Location/Qualifiers

FEATURES
source
1..667
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RH42996"
/sex="male and female"
/dev_stage="Adult"
/lab_host="DH5-alpha Tona"
/clone_lib="RH Drosophila melanogaster normalized Head
pFic-1"

BASE COUNT 196 a 184 c 182 g 105 t
Query Match 32.6%; Score 664.4; DB 12; Length 667;
Best Local Similarity 99.8%; Pred. No. 3.3e-76;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATTGTTGTCGACCGCTCGACGAAAGATATGAAATCGAGACGTAACGCAAAA 60
DB 2 ATTGTTGTCGACCGCTCGACGAAAGATATGAAATCGAGACGTAACGCAAAA 61
QY 61 CAGCTGGATTAAAGTTCCTTCGTAACGACGAGTGAAGAAAACAAAGTGAAGT 120
DB 62 CAGCTGGATTAAAGTTCCTTCGTAACGACGAGTGAAGAAAACAAAGTGAAGT 121
QY 121 AGGCAACAAAGCCGAATCAATCAATTTGAAAGCCAGTCAAGAGCAATCTTTA 180
DB 122 AGGCAACAAAGCCGAATCAATCAATTTGAAAGCCAGTCAAGAGCAATCTTTA 181
QY 181 ACCTCGCAAGTAATCCGTAACGACCGTGTAAACAAACGCAAGATTCTTCAAG 240
DB 182 ACCTCGCAAGTAATCCGTAACGACCGTGTAAACAAACGCAAGATTCTTCAAG 241
QY 241 AAAGCAATCATCTTCGAAATGAAGTCCCTGACGCGCTTCGCAAGAGTCCCT 300
DB 242 AAAGCAATCATCTTCGAAATGAAGTCCCTGACGCGCTTCGCAAGAGTCCCT 301
QY 301 ATGCGCGCTTGAATGCGAGCGGCGCATCGAGCGCCATCCAGCAACGCGGGA 360
DB 302 ATGCGCGCTTGAATGCGAGCGGCGCATCGAGCGCCATCCAGCAACGCGGGA 361
QY 361 GAGAACGCGGAAGTGAAGATGTATCTGTCCAACTGAAGAACCTGTTCCGTT 420
DB 362 GAGAACGCGGAAGTGAAGATGTATCTGTCCAACTGAAGAACCTGTTCCGTT 421
QY 421 AAAGACGAAAGCTCACCAAGCTGAGATCATTCAGACGTCATCATCTTGAC 480

```

Db 422 AAGAACAGAAAGCTCACCAAGCTGAGATCATCCAGACGTCATCGACTCATCTGCGAC 481
Qy 481 CTGACAGCCGAGCTGGAGACGACCCCGAATGGGCAATCTTGAATGGGACCGGCTGTG 540
Db 482 CTGACAGCCGAGCTGGAGACGACCCCGAATGGGCAATCTTGAATGGGACCGGCTGTG 541
Qy 541 ACCGCGGTGACGAGCTCCACGAGACGAGACGACGACATGAGATGCGGATGCCGAG 600
Db 542 ACCGCGGTGACGAGCTCCACGAGACGAGACGACGACATGAGATGCGGATGCCGAG 601
Qy 601 GCAGAACGGAAGTCGATCCAGATATCTCTGCCGCGCTGTAATGCCGAGCCGCGC 660
Db 602 GCAGAACGGAAGTCGATCCAGATATCTCTGCCGCGCTGTAATGCCGAGCCGCGC 661
Qy 661 AAGTC 666
Db 662 AAGTC 667

RESULT 7
A1134649 668 bp mRNA linear EST 19-APR-2001
LOCUS A1134649
DEFINITION GH12170.5, prime GH Drosophila melanogaster head port2 Drosophila melanogaster cDNA clone GH12170 5prime similar to M31900: D. melanogaster extramacrochaetae (emc) protein gene, exon 2, mRNA sequence.
ACCESSION A1134649
VERSION A1134649.1 GI:3627207
SOURCE EST.
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 668)
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S. and Rubin, G.M.
BDGP/HMI Drosophila EST Project
TITLE Unpublished
JOURNAL Contact: Stapleton, M.
COMMENT BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Plate: 121 row: F column: 10
High quality sequence stop: 555.
Location/Qualifiers
1. 668
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH12170"
/sex="male and female"
/dev stage="adult"
/lab_host="DHS - alpha"
/clone_lib="GH Drosophila melanogaster head port2"
/note="Organ: head; Vector: port2; Site: 1; EcorI; Site: 2; XhoI; Sized fractionated cDNAs were directly ligated into port2. Plasmid cDNA library."
BASE COUNT 254 a 126 c 104 g 184 t
ORIGIN

Query Match 32.5%; Score 661.6; DB 9; Length 668;
Best Local Similarity 99.4%; Pred. No. 7.5e-76;
Matches 664; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 962 GATCTCCGATTAAGTAACTGTAAGTCAAGAAACACAGCCGAGCGCATCTCT 1021
Db 1 GATCTCCGATTAAGTAACTGTAAGTCAAGAAACACAGCCGAGCGCATCTCT 60
Qy 1022 TCAACGCTCTTTCCCGGAAACGAAACCCGGAACGAGACTCGCTTTGCGCTC 1081
Db 1022 TCAACGCTCTTTCCCGGAAACGAAACCCGGAACGAGACTCGCTTTGCGCTC 1081

Db 61 TCAACGCTCTTTCCCGGAAACGAAACCCGGAACGAGACTCGCTTTGCGCTC 120
Qy 1082 CAACAGAGTTAGTTAAATTAATTAACAAGATACAAACAACAAACCTAGCATTA 1141
Db 121 CAACAGAGTTAGTTAAATTAATTAACAAGATACAAACAACAAACCTAGCATTA 180
Qy 1142 GACAGAGTTAGTTAAATTAATTAACAAGATACAAACAACAAACCTAGCATTA 1201
Db 181 GACAGAGTTAGTTAAATTAATTAACAAGATACAAACAACAAACCTAGCATTA 240
Qy 1202 CGCTTAAGATTAACGGAAGAAAGAAACCTGGCAAGCTTTAAACAATTAATTTCC 1261
Db 241 CGCTTAAGATTAACGGAAGAAAGAAACCTGGCAAGCTTTAAACAATTAATTTCC 300
Qy 1262 TTAACGCTTAATTAACGGAAGAAAGAAACCTGGCAAGCTTTAAACAATTAATTTCC 1321
Db 301 TTAACGCTTAATTAACGGAAGAAAGAAACCTGGCAAGCTTTAAACAATTAATTTCC 360
Qy 1322 GGTAAACAAACGCAAGAAAGAAAGTAACTAATGCAAGCAACAGATCAGCAATAA 1381
Db 361 GGTAAACAAACGCAAGAAAGAAAGTAACTAATGCAAGCAACAGATCAGCAATAA 420
Qy 1382 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1441
Db 421 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
Qy 1442 ATTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1501
Db 481 ATTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
Qy 1502 TTTCCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1561
Db 541 TTTCCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
Qy 1562 CACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1621
Db 601 CACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
Qy 1622 AGTTCATT 1629
Db 661 AGTTCATT 668

RESULT 8
BG636933 651 bp mRNA linear EST 23-APR-2001
LOCUS BG636933
DEFINITION SD14761.5, prime SD Drosophila melanogaster Schneider L2 cell culture port2 Drosophila melanogaster cDNA clone SD14761 5 similar to emc: Pband001007 'transcription factor' located on: 3L 61D1-61D2:: 04/13/2001, mRNA sequence.
ACCESSION BG636933
VERSION BG636933.1 GI:13764470
SOURCE EST.
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 651)
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S. and Rubin, G.M.
BDGP/HMI Drosophila EST Project
TITLE Unpublished
JOURNAL Contact: Stapleton, M.
COMMENT BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AB003469: arm:3L [608641, 911167]
estimated-cyto:61C8-61E1: 04/13/2001
Plate: SD.147 row: F column: 1
High quality sequence stop: 589.

FEATURES

source

Location/Qualifiers

1. 651
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="SD14/61"
 /lab_host="DHS-alpha"
 /clone_lib="SD Drosophila melanogaster Schneider L2 cell culture POT2"
 /note="Vector: POT2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into POT2. Plasmid cDNA library."
 BASE COUNT 194 a 181 c 177 g 99 t

Query Match 31.7%; Score 644.8; DB 10; Length 651;

Best Local Similarity 99.7%; Pred. No. 1.1e-73; Indels 0; Gaps 0;

Matches 646; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

15 CGCTCCGACGAAAGATATAGAAATCGAAGACGTAACGCAAAACACGCTGATTTAA 74
 1 CGTCCGACGAAAGATATAGAAATCGAAGACGTAACGCAAAACACGCTGATTTAA 60
 75 GTGTTTCTTTGCAACATCGACAGAGTGAAGAAACAAAGTGAAGTGAAGCCGACAAACCCG 134
 61 GTGTTTCTTTGCAACATCGACAGAGTGAAGAAACAAAGTGAAGTGAAGCCGACAAACCCG 120
 135 AATCAATCAATTTGAAAAAGCGATCAAGACATTTTAACTGAACTTCGCAAGTAAT 194
 121 AATCAATCAATTTGAAAAAGCGATCAAGACATTTTAACTGAACTTCGCAAGTAAT 180
 195 CCGGTACGACCGGTATCAACACGCGCAAGATTTCTACTCAAGAGCGTAAAGCATCATCTC 254
 181 CCGGTACGACCGGTATCAACACGCGCAAGATTTCTACTCAAGAGCGTAAAGCATCATCTC 240
 255 CAGAAATGAAGTCCCTGACCGGCGCTTCGACAGAGTGCCTTCGGAATGCCGCTTTGAA 314
 241 CAGAAATGAAGTCCCTGACCGGCGCTTCGACAGAGTGCCTTCGGAATGCCGCTTTGAA 300
 315 TGCACAGCGGCGCATCGACGCGCATCCGACGACCGCGGCGAGCGGAGCAACCCGAGAT 374
 301 TGCACAGCGGCGCATCGACGCGCATCCGACGACCGCGGCGAGCGGAGCAACCCGAGAT 360
 375 GAAGATGATATCTGCAAACTGAAGGACCTGCTTCATGATCCCAAGAACGAGAAAGCT 434
 361 GAAGATGATATCTGCAAACTGAAGGACCTGCTTCATGATCCCAAGAACGAGAAAGCT 420
 435 CACCAAGCTGAGATATCGACGACGTCATCGATCATCTGCAAGCTTCGACAGCCGAGCT 494
 421 CACCAAGCTGAGATATCGACGACGTCATCGATCATCTGCAAGCTTCGACAGCCGAGCT 480
 495 GAGAGCGACCCCGAGATGGGCACTTTCATGTCGGGACCGCTCTGAACGCGGTGAACG 554
 481 GAGAGCGACCCCGAGATGGGCACTTTCATGTCGGGACCGCTCTGAACGCGGTGAACG 540
 555 ACTCCACGAG 614
 541 ACTCCACGAG 600
 615 CGATCCAGATATCTGCGCCAGCGCTGAATGCCGAGAGCCGCGGAA 662
 601 CGATCCAGATATCTGCGCCAGCGCTGAATGCCGAGAGCCGCGGAA 648

RESULT 9

B1565419

631 bp mRNA linear EST 06-SEP-2001

LOCUS

B1565419

631 bp mRNA linear EST 06-SEP-2001

DEFINITION

B1565419

631 bp mRNA linear EST 06-SEP-2001

ACCESSION

B1565419

631 bp mRNA linear EST 06-SEP-2001

VERSION

B1565419

631 bp mRNA linear EST 06-SEP-2001

B1565419.1 GI:15454611
 B1565419.1 GI:15454611

KEYWORDS

EST.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 631)

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson

J., Champagne, M., Chavez, C., Doree, V., Farfan, D., Fries, E., George

R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S.,

Mungall, C.O., Nuno, J., Pacled, J., Paragas, V., Park, S.,

Phonamavong, S., Wan, K., Yu, C., Lewis, S.E., Ceiniker, S. and Rubin

G.M.

BDGP/HHMI RH Drosophila EST Project

Unpublished

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu

hit genomic AB003469; arm:3L (608641,91167)

estimated-cyto:61C8-61E1: 08/23/2001

Plate: RH.631 row: H column: 6

High quality sequence scop: 531.

Location/Qualifiers

1. 631

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="RH63190"

/sex="male and female"

/dev_stage="Adult"

/lab_host="DHS-alpha Tona"

/clone_lib="RH Drosophila melanogaster normalized Head

p1C-1"

/note="Organ: head; Vector: p1C1; Site 1: XhoI; Site 2:

BamHI; Library was kindly generated by Piero Carninci at

the RIKEN. The library was normalized and excised using

Cre recombinase. Plasmid cDNA library."

BASE COUNT 190 a 170 c 169 g 102 t

ORIGIN

Query Match 30.7%; Score 626; DB 12; Length 631;

Best Local Similarity 100.0%; Pred. No. 2.9e-71;

Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATTGTTGCTGACCGCTGCGACGAAAGATATAGAAATCGAAGACGTAACGCAAGAAA 60
 6 ATTGTTGCTGACCGCTGCGACGAAAGATATAGAAATCGAAGACGTAACGCAAGAAA 65
 61 CACGCTGATTAAGTGTTCCTTCGAAATCGACAGAGTGAAGAAACAAAGTGAAGTG 120
 66 CACGCTGATTAAGTGTTCCTTCGAAATCGACAGAGTGAAGAAACAAAGTGAAGTG 125
 121 ACGGCAAAAGCGGAAATCAATTAAGAAAGCAATCGAAGCAATCTTTAACTGA 180
 126 ACGGCAAAAGCGGAAATCAATTAAGAAAGCAATCGAAGCAATCTTTAACTGA 185
 181 ACTCGCAAGTATCCGATGACAGCGTGAACAACCGCAAGATTCTACTCAAGAGCGT 240
 186 ACTCGCAAGTATCCGATGACAGCGTGAACAACCGCAAGATTCTACTCAAGAGCGT 245
 241 AAGCAATCATCTCAAGATGAAGTCCGACGCGCTCTGCAAGAGAGTCCGCA 300
 246 AAGCAATCATCTCAAGATGAAGTCCGACGCGCTCTGCAAGAGAGTCCGCA 305
 301 ATCCGCGCTTGAATGCGGCGGCGCATCCAGCGCCATCCCAAGCAGCGGCGGCGG 360
 306 ATCCGCGCTTGAATGCGGCGGCGCATCCAGCGCCATCCCAAGCAGCGGCGGCGG 365
 361 GAGAAAGCGGCAAGATGAAGTATCTGTCCAACTGAAGACCTGCTTCGTTCAAGCC 420

Db 366 GAGAACCCGAGATGAATGTATCTGTCCAACTGAGAGACTCTTCCGTTTCATGCC 425

QY 421 AAGAACGAGAGCTCACCAGCTGAGATCATCCAGACGTCATCATCTGAGAC 480

Db 426 AAGAACGAGAGCTCACCAGCTGAGATCATCCAGACGTCATCATCTGAGAC 485

QY 481 CTGACGACCGAGCTGAGACGACCCGAGATGGGCACTTCGATGGGCGACCGCTCTG 540

Db 486 CTGACGACCGAGCTGAGACGACCCGAGATGGGCACTTCGATGGGCGACCGCTCTG 545

QY 541 ACCGCGGTGAACGAGCTCCACGAGAGAGAGACGACCATGATGCGGATGCCGAG 600

Db 546 ACCGCGGTGAACGAGCTCCACGAGAGAGAGAGACGACCATGATGCGGATGCCGAG 605

QY 601 GCAGAGCGGAGATGCATCCAGATAT 626

Db 606 GCAGAGCGGAGATGCATCCAGATAT 631

RESULT 10
BI227773 631 bp mRNA linear EST 11-JUN-2001

LOCUS BI227773
R224996, sprime RE Drosophila melanogaster normalized Embryo pFLC-1
Drosophila melanogaster cDNA clone R224996 5 similar to emc:
Pnan001007 'transcription factor' located on: 3L 61D1-61D2;
04/12/2001, mRNA sequence.

ACCESSION BI227773.1 GI:14695037

VERSION BI227773

KEYWORDS EST.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 631)

REFERENCE
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
, J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mista, S.,
Mungall, C.J., Nunoo, J., Pacleb, J., Pargass, V., Park, S.,
Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin
, G.M.
BDGP/HMT RE Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
hit genomic AB003469: arm:3L 1608641,9111671
estimated-cyto:61C8-61E1: 04/12/2001
Plate: RE:249 row: H column: 12
High quality sequence stop: 605.
Location/Qualifiers

FEATURES
Source

1..631
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="R224996"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha Tona"
/lab_host="RE Drosophila melanogaster normalized Embryo
pFLC-1"
/note="Organ: embryo; Vector: pFL1, Site_1: XhoI, Site_2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."

BASE COUNT 191 a 170 c 167 g 103 t

ORIGIN

Query Match 30.7%; Score 625.4; DB 12; Length 631;
Best Local Similarity 99.8%; Pred. No. 3.5e-71;

Matches 626; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTGTTGTGACCGCTCGACCAAAAGATATAGAAATGAGAAAGTAAACCAAGAAA 60

Db 5 ATTGTTGTGACCGCTCGACCAAAAGATATAGAAATGAGAAAGTAAACCAAGAAA 64

QY 61 CACGCTGATTTAAAGTGTGTTGTTGGAACATGACAGAGTGAAGAAACCAAGTGAAGT 120

Db 65 CACGCTGATTTAAAGTGTGTTGTTGGAACATGACAGAGTGAAGAAACCAAGTGAAGT 124

QY 121 AGGCGCAAAAGCCGGAATCAATCAATTTGAAAAGCAGTCAAGCAATCTTAACTGA 180

Db 125 AGGCGCAAAAGCCGGAATCAATCAATTTGAAAAGCAGTCAAGCAATCTTAACTGA 184

QY 181 ACCCTGCAATTAATCCGATGACCGGTGTAACACACCGCCAGATTCTAACAAGACGT 240

Db 185 ACCCTGCAATTAATCCGATGACCGGTGTAACACACCGCCAGATTCTAACAAGACGT 244

QY 241 AAGCAATCATCTCCAGAAATGAAGTCCCTGACGCGCTGTCAGACAGGTGCTCCGGA 300

Db 245 AAGCAATCATCTCCAGAAATGAAGTCCCTGACGCGCTGTCAGACAGGTGCTCCGGA 304

QY 301 ATGCCGCTTTGAATGCGACGCGCGCGATTCACCGCATTCACGACCGCGCGCGAG 360

Db 305 ATGCCGCTTTGAATGCGACGCGCGCGATTCACCGCATTCACGACCGCGCGAG 364

QY 361 GAGAACCGGAGATGAAGATGTATCTGTCCAACTGAGAGACTCTGTCATGATGCC 420

Db 365 GAGAACCGGAGATGAAGATGTATCTGTCCAACTGAGAGACTCTGTCATGATGCC 424

QY 421 AAGAACGAGAGCTCACCAGCTGAGATCATCCAGACGTCATCATCTGAGAC 480

Db 425 AAGAACGAGAGCTCACCAGCTGAGATCATCCAGACGTCATCATCTGAGAC 484

QY 481 CTGACGACCGAGCTGAGACGACCCGAGATGGGCACTTCGATGGGCGACCGCTCTG 540

Db 485 CTGACGACCGAGCTGAGACGACCCGAGATGGGCACTTCGATGGGCGACCGCTCTG 544

QY 541 ACCGCGGTGAACGAGCTCCACGAGAGAGAGAGACGACCATGATGCGGATGCCGAG 600

Db 545 ACCGCGGTGAACGAGCTCCACGAGAGAGAGAGACGACCATGATGCGGATGCCGAG 604

QY 601 GCAGAGCGGAGATGCATCCAGATATC 627

Db 605 GCAGAGCGGAGATGCATCCAGATATC 631

RESULT 11
BI232704 628 bp mRNA linear EST 11-JUN-2001

LOCUS BI232704
R228821, sprime RE Drosophila melanogaster normalized Embryo pFLC-1
Drosophila melanogaster cDNA clone R228821 5 similar to emc:
Pnan001007 'transcription factor' located on: 3L 61D1-61D2;
04/12/2001, mRNA sequence.

ACCESSION BI232704.1 GI:14700286

VERSION BI232704

KEYWORDS EST.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 628)

REFERENCE
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
, J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mista, S.,
Mungall, C.J., Nunoo, J., Pacleb, J., Pargass, V., Park, S.,
Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin
, G.M.
BDGP/HMT RE Drosophila EST Project
Unpublished
Contact: Stapleton, M.
BDGP

TITLE
JOURNAL
COMMENT

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AE003469: arm:3L [608641,91167]
estimated-cyto:61C8-61E1: 04/12/2001

Plate: RE.288 row: B column: 9
High quality sequence stop: 544.
Location/Qualifiers

FEATURES

source

1. 628
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE28821"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha toms"
/clone_lib="RE Drosophila melanogaster normalized Embryo pfic-1"
/note="Organ: embryo; Vector: pfic1; Site 1: XhoI; Site 2: BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cite recombinase. Plasmid cDNA library."

BASE COUNT 189 a 170 c 168 g 101 t

ORIGIN

Query Match 30.6%; Score 624; DB 12; Length 628;

Best Local Similarity 100.0%; Pred. No. 5.3e-71;

Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTCGTTGTCGACCGCTCCAGCAAAAGATATAGAAATCGAGAACTTAACGAGAAA 60
Db 5 ATTCTGTTGTCGACCGCTCCAGCAAAAGATATAGAAATCGAGAACTTAACGAGAAA 64

QY 61 CACGCTGATTAAGTGTTCGTTTCGACATCGACAGAGTGAAGAAACAAAGTGAAGT 120
Db 65 CACGCTGATTAAGTGTTCGTTTCGACATCGACAGAGTGAAGAAACAAAGTGAAGT 124

QY 121 AGGCAACAAAGCCGAATCAATCAATTTGAAAAAGCCAGTCAGAGCAATCTTAACTGA 180
Db 125 AGGCAACAAAGCCGAATCAATCAATTTGAAAAAGCCAGTCAGAGCAATCTTAACTGA 184

QY 181 ACCTCGAATTAATCCGGTACGACCGGTGAACAACGCCAAGATTCTACTCAAGAGCGT 240
Db 185 ACCTCGAATTAATCCGGTACGACCGGTGAACAACGCCAAGATTCTACTCAAGAGCGT 244

QY 241 AAAGCAATCATCTCCAGATGAAGTCCCTGACGCGCTCTGCAAGAGTGCCTCCGA 300
Db 245 AAAGCAATCATCTCCAGATGAAGTCCCTGACGCGCGCTCTGCAAGAGTGCCTCCGA 304

QY 301 ATGCCGCTTGAATGCGACGGGCGCATCCAGGCGCATCCGACGACCCGCGGACGGG 360
Db 305 ATGCCGCTTGAATGCGACGGGCGCATCCAGGCGCATCCGACGACCCGCGGACGGG 364

QY 361 GAGAACGCCGAGATGAAGATGTATCTGTCCAACTGAAGACCTCGTTCGTTCAATGCC 420
Db 365 GAGAACGCCGAGATGAAGATGTATCTGTCCAACTGAAGACCTCGTTCGTTCAATGCC 424

QY 421 AAAGCAAGAGTCAACCAAGTGAAGATCATCAGACGTCATCGATCATCTCGAC 480
Db 425 AAAGCAAGAGTCAACCAAGTGAAGATCATCAGACGTCATCGATCATCTCGAC 484

QY 481 CTGCAAGCCGAGCTCGAGACGACCCCGAGATGGGCACTTCGATGCGGACGCGCTCTG 540
Db 485 CTGCAAGCCGAGCTCGAGACGACCCCGAGATGGGCACTTCGATGCGGACGCGCTCTG 544

QY 541 ACCGCGGTGAACGAGCTCCAGAGACGAGACGAGACATGAGAGTGGAGTGGCCGAG 600
Db 545 ACCGCGGTGAACGAGCTCCAGAGACGAGACGAGACATGAGAGTGGAGTGGCCGAG 604

QY 601 GCAGAAAGCGAAGTGCATCCAGAT 624
Db 605 GCAGAAAGCGAAGTGCATCCAGAT 628

RESULT 12

BI637096

LOCUS

DEFINITION

BI637096 635 bp mRNA linear EST 10-SEP-2001
SD19014.5,prtime SD Drosophila melanogaster Schneider L2 cell culture
POT2 Drosophila melanogaster cDNA clone SD19014 5 similar to emc:
Pant001007 GO: [nucleus (GO:0005634); transcription co-repressor
(GO:0003714)] located on: 3L 61D1-61D2; 05/19/2001, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu

hit genomic AE003469: arm:3L [608641,91167]

estimated-cyto:61C8-61E1: 05/19/2001

Plate: SD.190 row: B column: 2

High quality sequence stop: 518.

Location/Qualifiers

1. 635

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="SD19014"

/lab_host="DHS-alpha"

/clone_lib="SD Drosophila melanogaster Schneider L2 cell

culture POT2"

/note="Vector: POT2; Site 1: EcoRI; Site 2: XhoI; Sized

fractionated cDNAs were directly ligated into pot2.

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Plasmid cDNA library."

Db 304 GCATCCAGCCCATCCACGACCGCGCGAGAAAGCCGAGATGATGTATC 363

Qy 386 TGTCCAAACTGAAGGAGCTGTGTTCCGTTCAATGCCCAAGAACAGGAGCTCAAGCTGG 445

Db 364 TGTCCAAACTGAAGGAGCTGTGTTCCGTTCAATGCCCAAGAACAGGAGCTCAAGCTGG 423

Qy 446 AGATCATCCAGCAGCTGATCATCTGATCTGCGACCTGACAGCCGAGCTGGAGACGAC 505

Db 424 AGATCATCCAGCAGCTGATCATCTGATCTGCGACCTGACAGCCGAGCTGGAGACGAC 483

Qy 506 CCGAGATGGCACTTGTGATGCGGACCGCTTGTGACGCGGTAAAGAGATCTCAAGAG 565

Db 484 CCGAGATGGCACTTGTGATGCGGACCGCTTGTGACGCGGTAAAGAGATCTCAAGAG 543

Qy 566 ACAGGACAGGACATGAGAGATGCGGATGCGGACGAGAAAGCGAAGTGCATCCAGTA 625

Db 544 ACAGGACAGGACATGAGAGATGCGGATGCGGACGAGAAAGTGCATCCAGTA 603

Qy 626 TCCTGCGCCAGCGCTGATGCGGACGCG 657

Db 604 TCCTGCGCCAGCGCTGATGCGGACGCG 635

RESULT 13

BI485228

LOCUS

DEFINITION 622 bp mRNA linear EST 28-AUG-2001

RE68523 5prime RE Drosophila melanogaster normalized Embryo pRc-1

Drosophila melanogaster cDNA clone RE68523 5 similar to emc:

Fan001007 GO:0005634; transcription co-repressor

(GO:0003141) located on: 3L 61D1-61D2; 05/16/2001, mRNA sequence.

BI485228

VERSION BI485228.1 GI:15325362

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 622)

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson

J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George

R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miera, S.,

Mungall, C.J., Nuno, J., Paclob, J., Paragas, V., Park, S.,

Pouanenvong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin

G.M.

BDGP/HM1 RE Drosophila EST Project

Unpublished

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_esc@fruitfly.berkeley.edu

hit genomic AB003469: arm:3L [608641,911167]

estimated-cyto:61C8-61E1: 05/16/2001

Plate: RE.685 row: B column: 11

High quality sequence stop: 601.

Location/Qualifiers

1..622

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="RE68523"

/sex="male and female"

/dev_stage="0-24 hours mixed stage embryonic"

/lab_host="DHS-alpha Tona"

/clone_lib="RE Drosophila melanogaster normalized Embryo

pRc-1"

/note="Organ: embryo; Vector: pRc1; Site 1: XhoI; Site 2:

BamHI; Library was kindly generated by Piero Carninci at

the RIKEN. The library was normalized and excised using

Cre recombinase. Plasmid cDNA library."

BASE COUNT 188 a 168 c 167 g 99 t

ORIGIN

Query Match 30.5%; Score 620.4; DB 12; Length 622;

Best Local Similarity 99.8%; Pred. No. 1.5e-70;

Matches 621; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATTGTTGTGACCCGTCGACGAAAGATATGAAATTCGAGAGCTTAAACGAGAAA 60

Db 1 ATTGTTGTGACCCGTCGACGAAAGATATGAAATTCGAGAGCTTAAACGAGAAA 60

Qy 61 CAGCTGGAATTAAGTTCGTTTGAACATGACAGAGTGAGAAAACAAAGTGAAGT 120

Db 61 CAGCTGGAATTAAGTTCGTTTGAACATGACAGAGTGAGAAAACAAAGTGAAGT 120

Qy 121 AGGCAACAAAGCCGAATCAATCAATTGAAAAGCAGTCAGAGCAATCTTTAACTGA 180

Db 121 AGGCAACAAAGCCGAATCAATCAATTGAAAAGCAGTCAGAGCAATCTTTAACTGA 180

Qy 181 ACTCGCAATTAATCCGATGCAACCGTTTAAACAGCGCAAGTTTACTCAAGAGCT 240

Db 181 ACTCGCAATTAATCCGATGCAACCGTTTAAACAGCGCAAGTTTACTCAAGAGCT 240

Qy 241 AAGCAATCATCTCCGAATGAAGTCCCTGACGCGCTGCGACAGGTGCTCCGGA 300

Db 241 AAGCAATCATCTCCGAATGAAGTCCCTGACGCGCTGCGACAGGTGCTCCGGA 300

Qy 301 ATGCGGCGCTTGAATGCGACGCGGCGCATCCAGCGCATCCACGACCGCGGAGCGG 360

Db 301 ATGCGGCGCTTGAATGCGACGCGGCGCATCCAGCGCATCCACGACCGCGGAGCGG 360

Qy 361 GAGAACGCGAGATGAAGATGATCTGTCCAACTGAAGACCTGTTCCGTTATGCC 420

Db 361 GAGAACGCGAGATGAAGATGATCTGTCCAACTGAAGACCTGTTCCGTTATGCC 420

Qy 421 AAGAACGAAAGCTCCACCAAGCTGAGATATCCAGAGCTCAATCATCTGCGAC 480

Db 421 AAGAACGAAAGCTCCACCAAGCTGAGATATCCAGAGCTCAATCATCTGCGAC 480

Qy 481 CTGCAAGCCAGCTGAGACGACCCGAGATGGGCAACTTCATGCGGACGCGCTCTG 540

Db 481 CTGCAAGCCAGCTGAGACGACCCGAGATGGGCAACTTCATGCGGACGCGCTCTG 540

Qy 541 ACAGCGGTGAACGACTCCACGAGACGACGACGACGACGACGACGACGACGACG 600

Db 541 ACAGCGGTGAACGACTCCACGAGACGACGACGACGACGACGACGACGACGACG 600

Qy 601 GCAGAACGGAAGTCGATCCAG 622

Db 601 GCAGAACGGAAGTCGATCCAG 622

RESULT 14

BI363964

LOCUS

DEFINITION 623 bp mRNA linear EST 01-AUG-2001

RD48965 5prime RE Drosophila melanogaster normalized Embryo pRc-1

Drosophila melanogaster cDNA clone RD48965 5 similar to emc:

Fan001007 'transcription factor' located on: 3L 61D1-61D2; 05/11/2001, mRNA sequence.

BI363964

VERSION BI363964.1 GI:15059992

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 623)

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson

J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George

R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miera, S.,

Mungall, C.J., Nuno, J., Paclob, J., Paragas, V., Park, S.,

Pouanenvong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin

G.M.

TITLE BDGP/HMI RE Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03469: arm:3L [608641,911167]
estimated-cyto:61C8-61E1: 05/13/2001
Plate: RE.489 row: F column: 5
High quality sequence stop: 530.
Location/Qualifiers

FEATURES

source

1. 623
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE48965"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo p1C-1"
/note="Organ: embryo; Vector: p1C1; Site 1: XhoI; Site 2: BamHI; Library was kindly generated by Pietro Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

BASE COUNT 187 a 169 c 167 g 100 t

ORIGIN

Query Match 30.4%; Score 619; DB 12; Length 623;
Best Local Similarity 100.0%; Pred. No. 2.3e-70;

Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTGTTGTCGACCGCTCGACCAAAAGATATGAAATGAGAACTTAACGAGAAA 60
DB 5 ATTGTTGTCGACCGCTCGACCAAAAGATATGAAATGAGAACTTAACGAGAAA 64
QY 61 CAGCTGATTAAGTGTTCGTTTTCGACATGACAGATGAGAAAACAAAGTGAAGT 120
DB 65 CAGCTGATTAAGTGTTCGTTTTCGACATGACAGATGAGAAAACAAAGTGAAGT 124
QY 121 AGGCAAAAGCCGAATCAATCAATGAAAAGCAGTCAAGAGCAATCTTTAATCTGA 180
DB 125 AGGCAAAAGCCGAATCAATCAATGAAAAGCAGTCAAGAGCAATCTTTAATCTGA 184
QY 181 ACCTCGAAGTAATCCGCTGACGCGTGAACAACCCCAAGATTCTCAAGAGCGT 240
DB 185 ACCTCGAAGTAATCCGCTGACGCGTGAACAACCCCAAGATTCTCAAGAGCGT 244
QY 241 AAGCAATCATCTCCAGATGAAGTCCCTGACGCGCTTCGACAGAGTCCCTCCGGA 300
DB 245 AAGCAATCATCTCCAGATGAAGTCCCTGACGCGCTTCGACAGAGTCCCTCCGGA 304
QY 301 ATGCGGCGCTTGAATGCAAGGCGCGCATCCAGCGCATCCAGCAACGCGCGGAGCG 360
DB 305 ATGCGGCGCTTGAATGCAAGGCGCGCATCCAGCGCATCCAGCAACGCGCGGAGCG 364
QY 361 GAGAAAGCCGAGATGATGTATCTGTCCAAACTGAAGAGCTGTTCCGTTCAAGCCC 420
DB 365 GAGAAAGCCGAGATGATGTATCTGTCCAAACTGAAGAGCTGTTCCGTTCAAGCCC 424
QY 421 AAGAAAGAGAGTCAACAAAGTGAAGATCATTCAGACGTCATGCAATCATCTGAC 480
DB 425 AAGAAAGAGAGTCAACAAAGTGAAGATCATTCAGACGTCATGCAATCATCTGAC 484
QY 481 CTGCAAGCGAGTGAAGAGCAACCCGAGATGAGGCACTTTCATGAGGAGCGCTCTG 540
DB 485 CTGCAAGCGAGTGAAGAGCAACCCGAGATGAGGCACTTTCATGAGGAGCGCTCTG 544
QY 541 ACCGCGTGAAGCGACTCCAGAGACGAGACGAGCATGAGAGATGCGAGTCCGAG 600
DB 545 ACCGCGTGAAGCGACTCCAGAGACGAGACGAGCATGAGAGATGCGAGTCCGAG 604

QY 601 GCAGAGCGAAGTCATC 619
DB 605 GCAGAGCGAAGTCATC 623

RESULT 15
BI173488
LOCUS
DEFINITION
BI173488 620 bp mRNA linear EST 09-JUL-2001
RE16419.5prime RE Drosophila melanogaster normalized Embryo p1C-1
Drosophila melanogaster cDNA clone RE16419 5 similar to emc:
Fban001007 'transcription factor' located on: 3L 61D1-61D2;:
04/12/2001, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BI173488.1 GI:114639295
EST.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 620)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miera, S.,
Mungall, C.J., Nuno, J., Pacle, J., Paragosa, V., Park, S.,
Phonemavong, S., Wan, K., Yo, C., Lewis, S.E., Celisner, S. and Rubin
G.M.

TITLE BDGP/HMI RE Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03469: arm:3L [608641,911167]
estimated-cyto:61C8-61E1: 04/12/2001
Plate: RE.164 row: B column: 7
High quality sequence stop: 530.
Location/Qualifiers

FEATURES

source

1. 620
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE16419"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DHS-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo p1C-1"
/note="Organ: embryo; Vector: p1C1; Site 1: XhoI; Site 2: BamHI; Library was kindly generated by Pietro Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

BASE COUNT 185 a 168 c 168 g 99 t

ORIGIN

Query Match 30.4%; Score 618.4; DB 12; Length 620;
Best Local Similarity 99.8%; Pred. No. 2.8e-70;

Matches 619; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATTGTTGTCGACCGCTCGACCAAAAGATATGAAATGAGAACTTAACGAGAAA 60
DB 1 ATTGTTGTCGACCGCTCGACCAAAAGATATGAAATGAGAACTTAACGAGAAA 60
QY 61 CAGCTGATTAAGTGTTCGTTTTCGACATGACAGATGAGAAAACAAAGTGAAGT 120
DB 61 CAGCTGATTAAGTGTTCGTTTTCGACATGACAGATGAGAAAACAAAGTGAAGT 120
QY 121 AGGCAAAAGCCGAATCAATCAATGAAAAGCAGTCAAGAGCAATCTTTAATCTGA 180
DB 121 AGGCAAAAGCCGAATCAATCAATGAAAAGCAGTCAAGAGCAATCTTTAATCTGA 180

```
Qy 181 ACCTCGCAAGTAATCCGTAAGACCGGTATACAAACCGCAAGATTCTACTCAAGAGCGT 240
    |||||
Db 181 ACCTCGCAAGTAATCCGTAAGACCGGTATACAAACCGCAAGATTCTACTCAAGAGCGT 240
Qy 241 AAAGCAATCATCTCCAGATGAAGTCCCTGACGGCCGTCTGCCAGACAGGTGCTCCGGA 300
    |||||
Db 241 AAAGCAATCATCTCCAGATGAAGTCCCTGACGGCCGTCTGCCAGACAGGTGCTCCGGA 300
Qy 301 ATGCCGGCCTTGAATGCGAGCGGGCGCATCCAGGGCCATCCACGCAACCGGGGAGAGGG 360
    |||||
Db 301 ATGCCGGCCTTGAATGCGAGCGGGCGCATCCAGGGCCATCCACGCAACCGGGGAGAGGG 360
Qy 361 GAGAACCCCGAGATGAAGATGTATCTGTCCAACTGAAGACCTGTTCCGTTCTATGCC 420
    |||||
Db 361 GAGAACCCCGAGATGAAGATGTATCTGTCCAACTGAAGACCTGTTCCGTTCTATGCC 420
Qy 421 AAGAACAGAAAGCTCACCAAGCTGGAATCATCCAGACGTCATCGACTATCATCTGCGAC 480
    |||||
Db 421 AAGAACAGAAAGCTCACCAAGCTGGAATCATCCAGACGTCATCGACTATCATCTGCGAC 480
Qy 481 CTGCGAGCCGAGCTGAGAGCGCACCCGAGATGGGCAACTTCGATGGGGGAGCCGCTCTG 540
    |||||
Db 481 CTGCGAGCCGAGCTGAGAGCGCACCCGAGATGGGCAACTTCGATGGGGGAGCCGCTCTG 540
Qy 541 ACGGCGGTGAACGGACTTCCACGAGAGAGGACAGCGACATGAGAGATGCCGATGCCGAG 600
    |||||
Db 541 ACGGCGGTGAACGGACTTCCACGAGAGAGGACAGCGACATGAGAGATGCCGATGCCGAG 600
Qy 601 GCAGAGCGGAAAGTCGATCC 620
    |||||
Db 601 GCAGAGCGGAAAGTCGATCC 620
```

Search completed: February 8, 2004, 10:22:19
Job time : 4781.14 secs